Clinical and genetic factors associated with disease course in inflammatory bowel disease

Spekhorst, Lieke Maaike

Clinical and genetic factors associated with disease course in inflammatory bowel disease

Lieke Maaike Spekhorst

L.M. Spekhorst

Clinical and genetic factors associated with disease course in inflammatory bowel disease

ISBN:	978-94-034-1106-4 (book)
	978-94-034-1105-7 (ebook)
Cover design:	Erwin Timmerman, Optima Grafische Communicatie
Layout:	Anna Bruijning, persoonlijkproefschrift.nl
Printed by:	Ipskamp Printing, proefschriften.net

This thesis was financially supported by University Medical Center of Groningen, Graduate School of Medical Sciences, the University of Groningen, Nederlandse vereniging voor Gastro-enterologie (NVGE), Takeda Nederland, Ferring B.V., Dr. Falk Pharma Benelux B.V., Mylan Healthcare B.V., Teva Netherlands B.V., Norgine B.V., Tramedico B.V. en Pfizer.

rijksuniversiteit groningen

Clinical and genetic factors associated with disease course in inflammatory bowel disease

Proefschrift

ter verkrijging van de graad van doctor aan de
Rijksuniversiteit Groningen
op gezag van de
rector magnificus prof. dr. E. Sterken
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

woensdag 14 november 2018 om 14.30 uur

door

Lieke Maaike Spekhorst

geboren op 24 augustus 1988
te Hengelo

Promotores

Prof. dr. R.K. Weersma

Prof. dr. G. Dijkstra

Copromotor

Dr. E.A.M. Festen

Beoordelingscommissie

Prof. dr. H.W. Verspaget

Prof. dr. H. Snieder

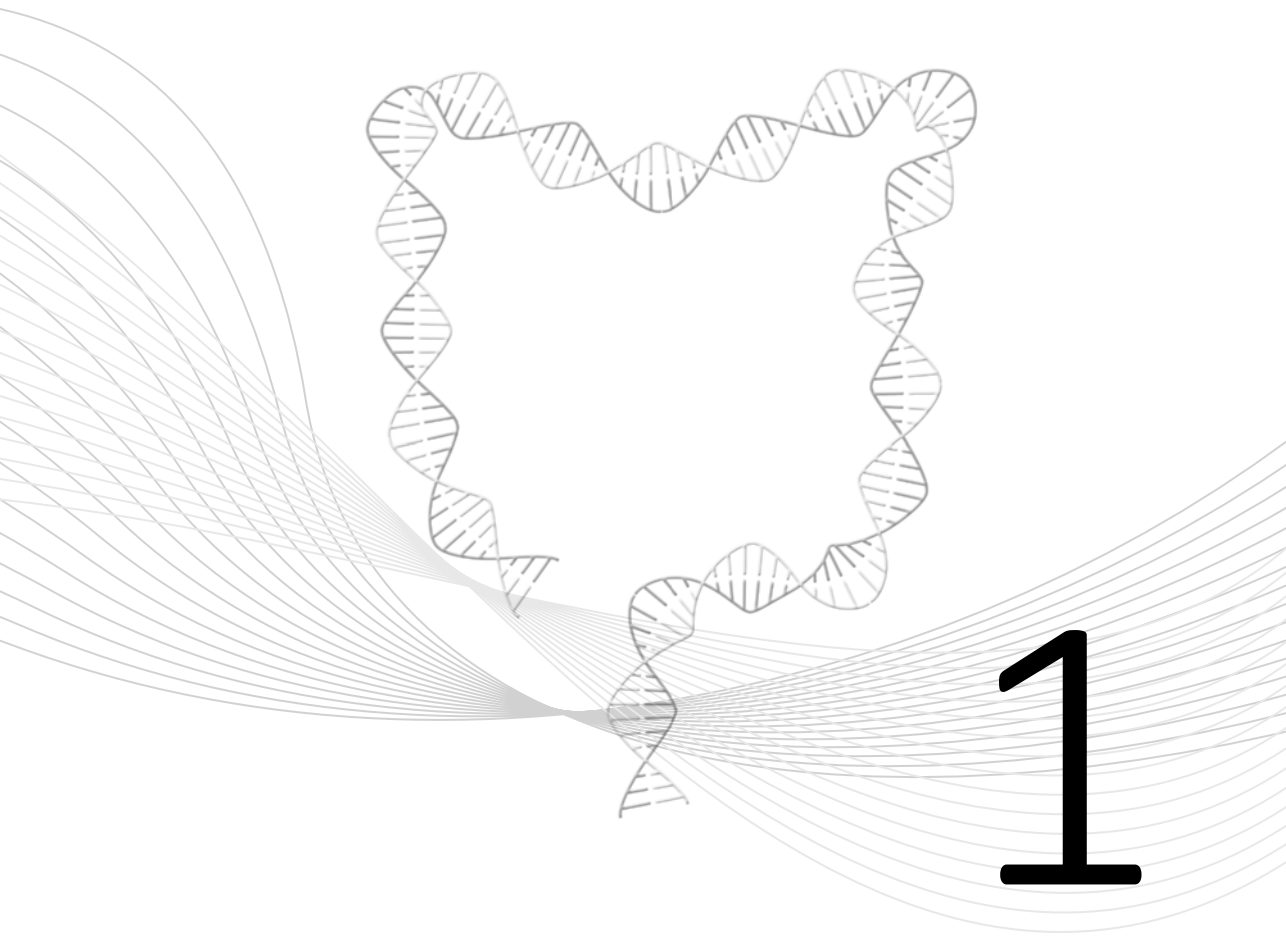
Prof. dr. G.R.A.M. D'Haens

Paranimfen

M.S. Spekhorst

L.S. Spekhorst

Table of contents



1

Introduction and outline of this thesis

Introduction and outline thesis

Inflammatory bowel disease (IBD) is a chronic immune-mediated inflammatory disease of the gastrointestinal tract that is characterized by its heterogeneous presentation and relapsing character. This highly variable presentation makes it very important to identify the clinical and genetic factors that can be used to predict disease course in individual patients. IBD comprises two main types: Crohn's disease (CD) and ulcerative colitis (UC).

Epidemiology

The prevalence of IBD varies greatly per geographic region, but also within these regions. Prevalence is highest in Europe and North America, with a prevalence of ~830 per 100,000 individuals in the Netherlands.[1] The high prevalence in Western nations is believed to be associated with their industrialization, and IBD prevalence is much lower in developing countries and Asia.[2] However, the incidence and prevalence of IBD in the Western world is now stabilizing, while the incidence of IBD in developing countries is increasing. Furthermore, people migrating from low prevalence areas to high prevalence areas become more susceptible to the disease, meaning that the children of adult immigrants acquire the same susceptibility to IBD as the native population, reaching the local prevalence of IBD within 1-2 generations. Together this evidence implies that factors in the Western lifestyle and environment play a role in IBD risk.[2,3]

Disease presentation of CD and UC

IBD patients are usually diagnosed in the second or third decade of life, and an early onset in patients with CD is correlated with a more severe disease phenotype.[4] IBD occurs more frequently in females, with a male:female ratio of 1:1.6 in CD and 1:1.2 in UC.[5] Beyond CD and UC there are several other subtypes of IBD. Patients with chronic colitis without clinical, endoscopic or histological features of either CD or UC are diagnosed with Inflammatory Bowel Disease Unclassified (IBD-U). The diagnosis Inflammatory Bowel Disease Indeterminate (IBD-I) is reserved for colitis patients in whom the pathologist does not find features pathognomonic for either CD or UC in the colectomy specimen.

Depending on disease localization and severity, patients with CD can suffer from symptoms such as abdominal pain, weight loss, fever and bloody diarrhoea. Inflammation in CD mainly affects the terminal ileum, but can occur anywhere in the digestive tract. It can also affect all mucosal layers, which can lead to strictures, abscesses and fistulas from the bowel to skin, bladder, vagina or other bowel segments. Surgical intervention can induce long-term remission[6] but is never a curative option in CD, therefore the principal treatment of inflammatory activity is drug therapy.[7]

Patients with UC typically have symptoms of bloody mucoid diarrhoea and lower abdominal pain. The inflammation in UC is confined to the colon, but in patients with a severe pancolitis a limited part of the terminal ileum can be affected, a condition called “backwash ileitis”. A total proctocolectomy will remove all affected tissue and is therefore curative in patients with UC. Colorectal cancer is a rare complication of IBD, and both CD and UC patients with longstanding inflammation of the colon have a higher risk for colorectal cancer compared to the general population.

Twenty-five percent of patients with IBD have extraintestinal manifestations (EIMs).[8] The most common EIMs are joint complaints (sacroiliitis and ankylosing spondylitis), ocular involvement (uveitis and episcleritis), and skin involvement (pyoderma gangrenosum, erythema nodosum, psoriasis or palmoplantar psoriasiform pustolosis, and metastatic CD). The prevalence of EIMs is higher in CD compared to UC,[9] and active smoking increases risk for EIMs in CD and UC.[10] Primary sclerosing cholangitis (PSC) is a rare liver disease characterized by fibrosis of the bile duct that often co-occurs with IBD. Approximately 68% of patients with PSC have IBD, of which 98% have UC or colonic CD and 6% have isolated ileal CD.[11] Increased risk of thromboembolic events and osteoporosis are considered complications of IBD as opposed to true EIMs. An increased risk of thromboembolic events is associated with disease flares.[12,13] The factors associated with osteoporosis are sustained use of corticosteroids and malabsorption.

Treatment strategies in IBD consist of surgical intervention and medication. Selection of appropriate treatment depends on disease behaviour, location and severity. Choice of medication is further influenced by medication potency, potential side effects and the presence of EIMs or complications.[14,15] Medications used in IBD consist of corticosteroids, mesalazine, immunomodulators (thiopurines or methotrexate) and biologics (anti-TNFα, anti-IL12/IL23 and anti-$\alpha_4\beta_7$ agents). However, 16% of UC patients and up to 47% of CD patients will still require surgery in the ten years following their diagnosis.[16] CD patients with disease activity are often initially treated with corticosteroids, and patients with steroid-refractory disease or steroid dependent disease are often treated with immunomodulators. Anti-TNFα therapy is indicated in patients with severe or refractory disease in whom immunosuppressive drugs are failing. Some patients who use anti-TNFα agents develop anti-TNFα antibodies, which reduces the anti-inflammatory effects of the drug.[17]

UC patients with proctitis can often be treated with local therapy, but when this is insufficient, or the inflammation affects a larger extent of the colon, mesalazine should be started, sometimes in combination with corticosteroids. Immunomodulators can be added later to prevent relapse of inflammation.

Pathogenesis of IBD

Inflammation arises in IBD through an exacerbated immune response to the commensal bacteria in the gut in genetically predisposed individuals. While this process may sound straightforward, environmental factors,[18] the immune response, the gut microbiota and genetics all play a role in disease risk, making the study of the pathogenesis of IBD a complex affair.

Smoking and appendectomy are environmental factors well established to play a role in IBD risk. Being a former smoker is a risk factor for the development of UC, while smoking is protective of an exacerbation once the disease has established. On the other hand, being a current smoker is a risk factor for the development of CD and is associated with a higher risk of disease exacerbations.[19] Prior appendectomy is an environmental factor associated with a reduced risk of acquiring UC.[20]

In recent years, the gut microbiome has become the focus of increased attention for its role in IBD. Many studies report dysbiosis, showing a decreased diversity in the gut microbiome in IBD patients as compared to that of healthy controls. In animal models of gut inflammation it has been shown that some microbial species reduce intestinal inflammation, whereas others (some known to be correlated with IBD) can exacerbate it, implying an important role for the gut microbiome in inflammation in IBD.[21-25]

The genetic background of IBD risk has been studied extensively in the recent and rapid growth of the field of genetic research. Genome-wide association studies (GWAS) initially revealed more than 99 susceptible loci associated with IBD risk.[26] In GWAS, large numbers of patients and healthy individuals are genotyped and compared for thousands of small common genetic variants, called single nucleotide polymorphisms (SNPs). As GWAS for other immune-mediated disease were carried out, it became clear that most genetic risk variants were shared between IBD and other immune-mediated diseases.[27,28] With this in mind, the Immunochip was designed. It is a customized array covering ~200,000 SNPs selected from GWAS results for 12 immune-mediated diseases. Immunochip studies have now been very successful in identifying new genetic risk loci for IBD, increasing the number of known IBD risk loci to 242.[29-31]

Predicting disease course using clinical and genetic risk factors

A major concern in the management and treatment of IBD is the heterogeneity of the disease between patients. The localization, severity and course of the disease can be highly variable between individuals with the same diagnosis, yet we still cannot predict which patients are at risk of a severe disease course and who may require surgical intervention and expensive medication such as biologicals. It is thus important to identify clinical parameters that can predict disease course.

Clinical parameters associated with a severe disease course, such as early age of onset, current smoking, and fistulising disease, should be considered as predictors for severe disease in clinical practice. These factors could, for example, be a reason for starting anti-TNF treatment or performing

surgical intervention earlier in disease course. Smoking, another clinical parameter associated with severe disease course, should be targeted by clinical support for smoking cessation before increasing the intensity of medical treatment. In addition to these clinical parameters known to be associated with severe disease course, there are many factors whose effect on disease phenotype and disease course has not yet been intensively researched, including the influence of ethnicity and sex differences.

Genetic variants should also be considered in efforts to predict disease course. When it comes to disease risk, the role of genetics has been studied extensively. However, studies on genetic risk loci associated with a particular IBD phenotype remain few. A large international collaboration found three risk loci associated with disease location, but little or no genetic association with disease behaviour.[32] Genetic studies on disease course and response to therapy have only just started to evolve and there is much more information to gain in this field of research.

Ultimately, we want to incorporate both clinical and genetic risk factors into an algorithm that can predict whether a given patient will be at risk of a severe disease course. With such a prediction model, we could prevent unfavourable outcomes by starting harsher medical treatment earlier in the disease course in patients at high risk for severe disease outcome. The aim of my thesis is therefore to identify clinical and genetic predictors for specific IBD phenotypes.

Outline of this thesis

This thesis focuses on resolving important questions in two areas of interest in IBD disease course:

- The epidemiology of IBD disease course: Identifying clinical factors that explain phenotypic differences in IBD patients
- The genetics of IBD disease course: Identifying genetic risk loci associated with IBD disease behaviour

Cohorts used in this thesis

To identify clinical and genetic risk factors for specific IBD phenotypes, it is crucial to have detailed data that has been collected in a standardized manner. In the first part of my thesis I use data made available through the Parelsnoer Institute (www.parelsnoer.org). At the time of my study, the IBD Parelsnoer cohort contained 3388 patients with IBD collected via a collaboration of the eight University Medical Centers (UMCs) in the Netherlands. Detailed phenotypic data was collected for 225 IBD-related items. In the second part of my thesis I use the Understanding and Redefining IBD (UR-IBD) cohort collected through the department of Gastroenterology and Hepatology in the University Medical Center Groningen (UMCG). The UR-IBD cohort currently consists of over 1000 patients with IBD, for whom we have extensive phenotype, genotype, and microbiome data as well as serological markers.

Large cohorts are required to assess correlations between clinical factors and IBD phenotypes and to identify genetic risk loci associated with specific IBD phenotypes. Therefore, I collaborated with the University of Utrecht, where a large IBD cohort of 2252 IBD patients, the COIN study (Costs Of Inflammatory bowel disease in the Netherlands), had been collected from both the University and general hospitals.[33] I used this cohort to increase my sample size to gain more power to assess phenotypic differences in patients with IBD. I also collaborated with the Translational Research in Gastrointestinal Disorders (TARGID) research center of the University of Leuven, whose cohort has both genotype data and infliximab antibody data available. Patients in the IBD Parelsnoer cohort, the UR-IBD cohort and the Leuven cohort have all been genotyped using the Immunochip.

Part I. The epidemiology of IBD disease course: Identifying clinical factors that explain phenotypic differences in IBD patients

To identify factors associated with a specific phenotype, it is critical that patient phenotypes are described in a consistent manner. The Montreal classification is a classification system for subphenotypes of both CD and UC. While Montreal classification is used extensively in the clinic, there is virtually no data on its reliability and reproducibility. Therefore, in **chapter 2**, we validated the Montreal classification among 30 observers with different professions (gastroenterologist specialist in IBD, gastroenterologist in training and IBD-nurses) in 20 de-identified medical records. This intra- and inter-observer variability validation is important because **chapter 3** gives an overview of the phenotypic characteristics of patients with IBD present in the IBD Parelsnoer cohort on 17 July 2014. Furthermore, we used the IBD Parelsnoer cohort to further explore the clinical parameters that are associated with phenotypic differences in patients with IBD.

As IBD is most often diagnosed in the second or third decade of life, patients can experience difficulties at work or study early in life. IBD can be a significant burden in daily life, as patients suffer from disease symptoms or flares of the disease. Consequently, patients with IBD are at significant risk for work disability. In **chapter 4** we assess potential risk factors for work disability in IBD, which is of great importance in efforts to improve IBD patient care.

IBD is emerging as a global disease, with a considerable variation within and between geographic regions. The incidence of IBD in developing countries may be increasing because of industrialization and Western lifestyle changes, but these changes cannot explain the phenotypic heterogeneity within and between regions. Population-based studies concerning phenotype differences between ethnicities and different countries of birth remain scarce. In **chapter 5** we explore the role of ethnicity and country of birth on IBD phenotype.

In recent years, diagnostic and treatment strategies for most chronic diseases have increasingly been adjusted to the individual patient, with extra care being taken to address the differential needs of male versus female patients, for example in cardiovascular disease.[34] In IBD care, both

diagnostic and treatment strategies are applied equally to male and female patients. If we want to move further towards personalized treatment, it is important to assess differences in IBD disease course and in disease phenotypes between sexes. **Chapter 6** therefore focuses on differences between male and female IBD patients, comparing phenotype, clinical manifestations, disease course, medical treatment and other healthcare consumption. For this study we used both the IBD Parelsnoer cohort and the COIN study cohort.

Part II. The genetics of IBD disease course: Identifying genetic risk loci that are associated with IBD disease behaviour

Although more than 200 IBD loci are known to be associated with IBD risk there are only a few studies that correlate these loci to IBD disease behaviour. In this second part of my thesis I focus on IBD-associated genetic variants that are associated with a particular disease phenotype.

Chapter 7 is a review of IBD genetics. It describes the clinical presentation of IBD and gives an overview of the progress that has been made in the field of IBD genetics, from linkage and candidate studies to GWAS and Immunochip studies. Chapter 7 also describes the genetic and biological pathways of IBD and its overlap with other immune-mediated disease. The final part of this review focuses on genetic findings and how we can translate them to clinical practice.

Hidradenitis suppurativa (HS) is a chronic inflammation of the apocrine glands often followed by sinus tract formation and scarring. Although HS is not yet a well-recognized EIM in IBD, the prevalence of HS in IBD is much higher than in the general population, which raises the hypothesis of a shared pathogenesis. In **chapter 8**, I aimed to identify genetic and clinical parameters associated with the occurrence of HS in IBD.

The need for surgery for fibrostenotic disease in patients with CD is an indicator of a severe disease course. In **chapter 9**, I aimed to identify disease-modifying genes for recurrent fibrostenotic disease behaviour in CD. I used Immunochip data to perform a within-cases analysis in two independent cohorts by comparing patients with fibrostenotic CD with patients with purely inflammatory CD.

The use of anti-TNFα agents is very important for inducing and maintaining clinical remission in patients with CD. However, some patients develop anti-drug antibodies (ADAs) to anti-TNFα, resulting in loss of response. In **chapter 10**, I aimed to identify genetic variants that play a role in the development of ADAs to anti-TNFα (infliximab and adalimumab), comparing IBD patients who developed anti-TNFα ADAs to IBD patients without these ADAs. This study was carried out in collaboration with University of Leuven, Belgium.

Chapter 11 gives an overview of the studies present in this thesis, discussion and the future perspectives.

References

1. van den Heuvel TRA, Jeuring SFG, Zeegers MP, *et al.* A 20-Year Temporal Change Analysis in Incidence, Presenting Phenotype and Mortality, in the Dutch IBDSL Cohort-Can Diagnostic Factors Explain the Increase in IBD Incidence? *J Crohn's Colitis* 2017;11:1169-79.
2. Logan I, Bowlus CL. The geoepidemiology of autoimmune intestinal diseases. *Autoimmun Rev* 2010;9:A372-8.
3. Molodecky NA, Soon IS, Rabi DM, *et al.* Increasing Incidence and Prevalence of the Inflammatory Bowel Diseases With Time, Based on Systematic Review. *Gastroenterology* 2012;142:46-54.e42.
4. Van Limbergen J, Russell RK, Drummond HE, *et al.* Definition of phenotypic characteristics of childhood-onset inflammatory bowel disease. *Gastroenterology* 2008;135:1114-22.
5. Wagtmans MJ, Verspaget HW, Lamers CB, *et al.* Gender-related differences in the clinical course of Crohn's disease. *Am J Gastroenterol* 2001;96:1541-6.
6. Ponsioen CY, de Groof EJ, Eshuis EJ, *et al.* Laparoscopic ileocaecal resection versus infliximab for terminal ileitis in Crohn's disease: a randomised controlled, open-label, multicentre trial. *Lancet Gastroenterol Hepatol* 2017;2:785-92.
7. Travis SPL, Stange EF, Lémann M, *et al.* European evidence based consensus on the diagnosis and management of Crohn's disease: current management. *Gut* 2006;55:i16-35.
8. Rankin GB, Watts HD, Melnyk CS, *et al.* National Cooperative Crohn's Disease Study: extraintestinal manifestations and perianal complications. *Gastroenterology* 1979;77:914-20.
9. Lakatos PL, Lakatos L, Kiss LS, *et al.* Treatment of extraintestinal manifestations in inflammatory bowel disease. *Digestion* 2012;86 Suppl 1:28-35.
10. Severs M, van Erp SJH, van der Valk ME, *et al.* Smoking is Associated With Extra-intestinal Manifestations in Inflammatory Bowel Disease. *J Crohn's Colitis* 2016;10:455-61.
11. Boonstra K, Weersma RK, van Erpecum KJ, *et al.* Population-based epidemiology, malignancy risk, and outcome of primary sclerosing cholangitis. *Hepatology* 2013;58:2045-55.
12. Kappelman MD, Horvath-Puho E, Sandler RS, *et al.* Thromboembolic risk among Danish children and adults with inflammatory bowel diseases: a population-based nationwide study. *Gut* 2011;60:937-43.
13. Grainge MJ, West J, Card TR. Venous thromboembolism during active disease and remission in inflammatory bowel disease: a cohort study. *Lancet* 2010;375:657-63.
14. Dignass A, Lindsay JO, Sturm A, *et al.* Second European evidence-based consensus on the diagnosis and management of ulcerative colitis Part 2: Current management. *J Crohn's Colitis* 2012;6:991-1030.
15. Dignass A, Van Assche G, Lindsay JO, *et al.* The second European evidence-based Consensus on the diagnosis and management of Crohn's disease: Current management. *J Crohn's Colitis* 2010;4:28-62.
16. Frolkis AD, Dykeman J, Negrón ME, *et al.* Risk of surgery for inflammatory bowel diseases has decreased over time: a systematic review and meta-analysis of population-based studies. *Gastroenterology* 2013;145:996-1006.
17. Allez M, Karmiris K, Louis E, *et al.* Report of the ECCO pathogenesis workshop on anti-TNF therapy failures in inflammatory bowel diseases: definitions, frequency and pharmacological aspects. *J Crohns Colitis* 2010;4:355-66.
18. van der Sloot KWJ, Amini M, Peters V, *et al.* Inflammatory Bowel Diseases. *Inflamm Bowel Dis* 2017;23:1499-509.
19. van der Heide F, Dijkstra A, Weersma RK, *et al.* Effects of active and passive smoking on disease course of Crohn's disease and ulcerative colitis. *Inflamm Bowel Dis* 2009;15:1199-207.
20. Molodecky NA, Kaplan GG. Environmental risk factors for inflammatory bowel disease. *Gastroenterol Hepatol (N Y)* 2010;6:339-46.
21. Sartor RB. Microbial Influences in Inflammatory Bowel Diseases. *Gastroenterology* 2008;134:577-94.
22. Kostic AD, Xavier RJ, Gevers D. The microbiome in inflammatory bowel disease: current status and the future ahead. *Gastroenterology* 2014;146:1489-99.

23. Atarashi K, Tanoue T, Oshima K, *et al.* Treg induction by a rationally selected mixture of Clostridia strains from the human microbiota. *Nature* 2013;500:232-6.

24. Morgan XC, Tickle TL, Sokol H, *et al.* Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol* 2012;13:R79.

25. Imhann F, Vich Vila A, Bonder MJ, *et al.* Interplay of host genetics and gut microbiota underlying the onset and clinical presentation of inflammatory bowel disease. *Gut* 2018;67(1):108-119.

26. Anderson CA, Boucher G, Lees CW, *et al.* Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet* 2011;43:246-52.

27. Zhernakova A, van Diemen CC, Wijmenga C. Detecting shared pathogenesis from the shared genetics of immune-related diseases. *Nat Rev Genet* 2009;10:43-55.

28. Uniken Venema WT, Voskuil MD, Dijkstra G, *et al.* The genetic background of inflammatory bowel disease: from correlation to causality. *J Pathol* 2017;241:146-58.

29. de Lange KM, Moutsianas L, Lee JC, *et al.* Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat Genet* 2017;49:256-61.

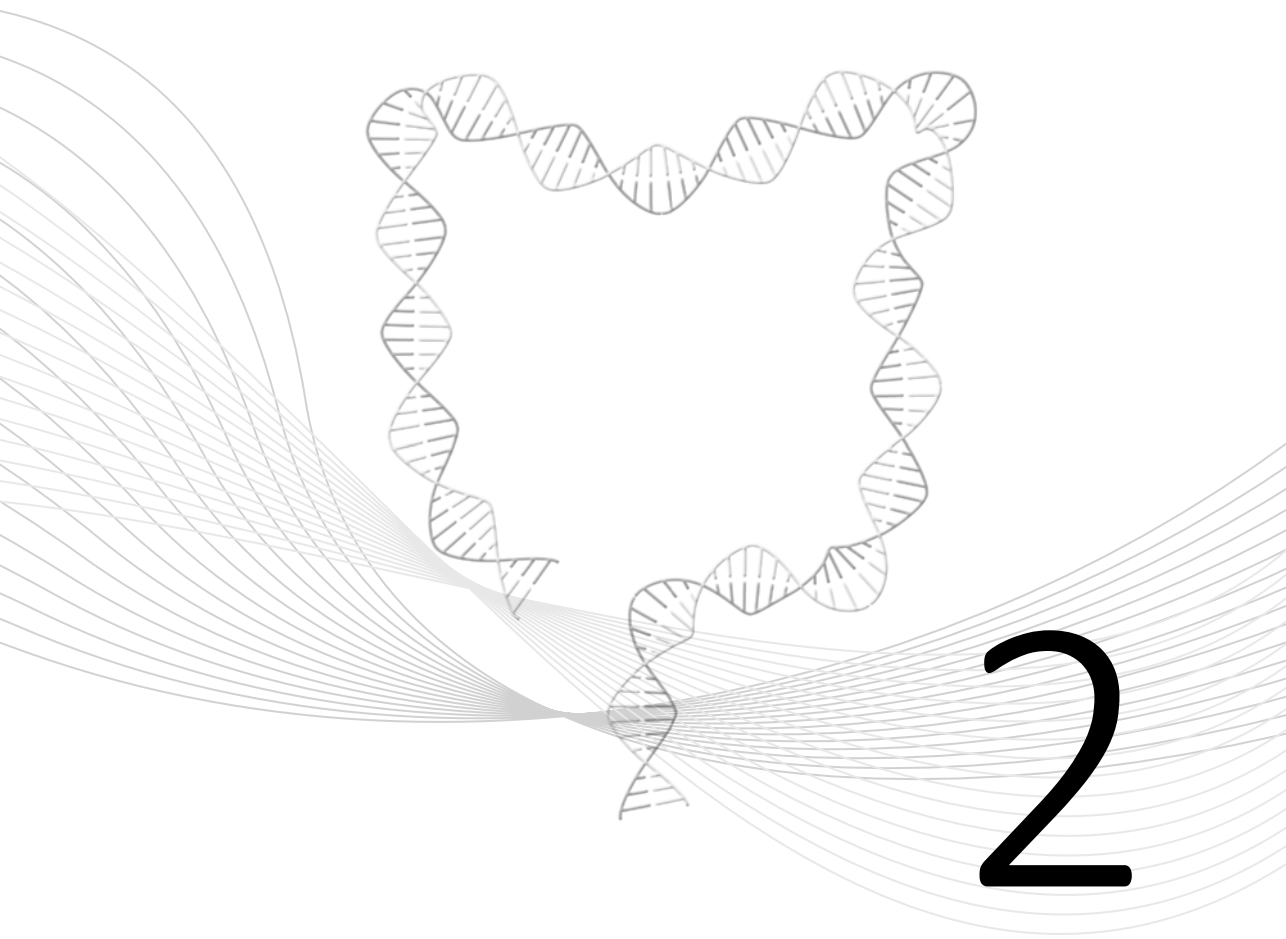
30. Liu JZ, van Sommeren S, Huang H, *et al.* Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet* 2015;47:979-86.

31. Jostins L, Ripke S, Weersma RK, *et al.* Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;491:119-24.

32. Cleynen I, Boucher G, Jostins L, *et al.* Inherited determinants of Crohn's disease and ulcerative colitis phenotypes: a genetic association study. *Lancet* 2016;387:156-67.

33. van der Valk ME, Mangen M-JJ, Leenders M, *et al.* Healthcare costs of inflammatory bowel disease have shifted from hospitalisation and surgery towards anti-TNFα therapy: results from the COIN study. *Gut* 2014;63:72-9.

34. Khamis RY, Ammari T, Mikhail GW. Gender differences in coronary heart disease. *Heart* 2016;102:1142-9.



2

Performance of the Montreal classification for inflammatory bowel diseases

LM Spekhorst*, MC Visschedijk*, R Alberts, EA Festen,
E-J van der Wouden, G Dijkstra, RK Weersma;
on behalf of the Dutch Initiative on Crohn and Colitis (ICC)
*Authors contributed equally

World J Gastroenterol 2014 Nov 7;20(41):15374-15381

Abstract

Background: To validate the Montreal classification system for Crohn's disease (CD) and ulcerative colitis (UC) within the Netherlands.

Methods: A selection of 20 de-identified medical records with an appropriate representation of the inflammatory bowel disease (IBD) subphenotypes were scored by 30 observers with different professions (gastroenterologist specialist in IBD, gastroenterologist in training and IBD-nurses) and experience level with IBD patient care. Patients were classified according to the Montreal classification. In addition, participants were asked to score extraintestinal manifestations (EIM) and disease severity in CD based on their clinical judgment. The inter-observer agreement was calculated by percentages of correct answers (answers identical to the "expert evaluation") and Fleiss-kappa (κ). Kappa cut-offs: < 0.4-poor; 0.41-0.6-moderate; 0.61-0.8-good; > 0.8 excellent.

Results: The inter-observer agreement was excellent for diagnosis (κ = 0.96), perianal disease (κ = 0.92) and disease location in CD (κ = 0.82) and good for age of onset (κ = 0.67), upper gastrointestinal disease (κ = 0.62), disease behaviour in CD (κ = 0.79) and disease extent in UC (κ = 0.65). Disease severity in UC was scored poor (κ = 0.23). The additional items resulted in a good inter-observer agreement for EIM (κ = 0.68) and a moderate agreement for disease severity in CD (κ = 0.44). Percentages of correct answers over all Montreal items give a good reflection of the inter-observer agreement (> 80%), except for disease severity (48%-74%). IBD-nurses were significantly worse in scoring upper gastrointestinal disease in CD compared to gastroenterologists (P = 0.008) and gastroenterologists in training (P = 0.040). Observers with less than 10 years of experience were significantly better at scoring UC severity than observers with 10-20 years (P = 0.003) and more than 20 years (P = 0.003) of experience with IBD patient care. Observers with 10-20 years of experience with IBD patient care were significantly better at scoring upper gastrointestinal disease in CD than observers with less than 10 years (P = 0.007) and more than 20 years (P = 0.007) of experience with IBD patient care.

Conclusion: We found a good to excellent inter-observer agreement for all Montreal items except for disease severity in UC (poor).

Introduction

Inflammatory bowel diseases (IBD) are common, chronic relapsing gastrointestinal inflammatory disorders consisting of mainly two diseases: Crohn's disease (CD) and ulcerative colitis (UC). IBD affects approximately 1 in 1000 individuals in Western Europe.[1,2]

In CD inflammation is transmural and can occur throughout the entire gastrointestinal tract, in UC the inflammation is limited to the mucosal layer of the colon.[3,4] In addition to intestinal inflammation, up to 25% of the patients have extraintestinal symptoms like uveitis, arthritis and erythema nodosum. Management of IBD with drug therapy consists of mesalazine, corticosteroids, and immunosuppressants like azathioprine and anti-tumour necrosis factor (TNF) antibody therapies. Most of these treatments have significant side effects, are expensive and often ineffective. Half of the patients (25%-30% in UC, 70%-75% in CD) require surgical intestinal resections because of refractory disease, fibrostenotic disease, abscesses, fistulae or the development of colorectal cancer.[5-9]

The pathogenesis of IBD is still not fully understood. The current hypothesis is that it arises from an inappropriate activation of the mucosal immune system in response to commensal bacteria in a genetically susceptible host.[10,11] Several biological pathways that play a role in this inappropriate inflammation have been identified through genetic studies. Recently, the International IBD Genetics Consortium has identified 163 independent genetic susceptibility loci.[12-15] However, the translation of biological knowledge on the pathogenesis of IBD towards the clinic is complicated by the great variety in the clinical presentation of IBD. For both clinical and genetic research it is of great importance that phenotypes of patients are described in a consistent manner.

In 2000 the Vienna classification was introduced, which was the first attempt to classify different clinical phenotypes of CD.[16] The Vienna classification was followed by the Montreal classification in 2008.[17] The Montreal classification describes the extent and behaviour of CD in more detail and includes a classification system for UC (Table 1).[17] Although the Montreal classification is widely used in both research and clinical practice, there is very limited data available on its reliability. Only two studies assessed the inter-observer reliability and validity of the Montreal classification, an Australian-New Zealand study and a study of the National Institutes of Diabetes and Digestive and Kidney Diseases IBD Genetics Consortium. Both studies had a small number of observers. The Australian-New Zealand study assessed only reliability of the Montreal classification in CD. In both studies the inter-observer agreement was good for disease location, but only moderate/fair for upper gastrointestinal involvement.[18,19]

The aim of this study is to validate the Montreal phenotype classification for both CD and UC in the Netherlands. Secondly, we will assess the influence of one's profession (gastroenterologist,

gastroenterologist in training, IBD-nurse) and level of experience (< 10 years, 10-20 years, > 20 years) on the reliability of the Montreal classification scoring.

Table 1. Montreal classification of Crohn's disease, ulcerative colitis, non-classified chronic colitis and indeterminate colitis

Diagnosis (20 case-vignettes)	
Crohn's Disease (CD) Ulcerative colitis (UC) Non-classified chronic colitis (IBD-U) Indeterminate colitis (IBD-I)	
Age of onset (A) (20 case-vignettes)	
A1: 16 years or younger A2: 17-40 years A3: over 40 years	
CD (10 case-vignettes)	UC, IBD-U, IBD-I (10 case-vignettes)
Localization (L)	Disease extent (E)
L1: Terminal ileum L2: Colon L3: Ileocolon L4: Upper gastrointestinal P: Perianal disease	E1: Proctitis E2: Left-sided UC; proximal extent of inflammation is distal to the rectosigmoid E3: Extensive UC; involvement extends proximal to the splenic flexure.
Behaviour (B)	Disease severity (S)
B1: Nonstricturing, nonpenetrating B2: Stricturing B3: Penetrating	S0: Remission, no symptoms S1: Mild, ≤ 4x/day stools, no systemic signs of toxicity, normal ESR S2: Moderate, > 4x/day stools, minimal systemic signs of toxicity S3: Severe, ≥ 6x/day stools, pulse > 90 beats/min, temperature > 37,5, Haemoglobin < 6,5 mmol/L, ESR > 30mm

EIM: Extraintestinal manifestations; ESR: Erythrocyte sedimentation rate.

Methods

Cases and observers

Twenty patient records were selected from the specialized IBD unit of the Department of Gastroenterology and Hepatology of the University Medical Center Groningen, the Netherlands (10 case vignettes) and the IBD unit of the Gastroenterology and Hepatology department of the non-university medical center Isala Clinics, Zwolle, the Netherlands (10 case vignettes). The case vignettes consisted of clinical-, endoscopy-, pathology- and operation reports. All case vignettes

were anonymised and the selection gave an appropriate representation of the IBD subphenotypes (Figure 1).

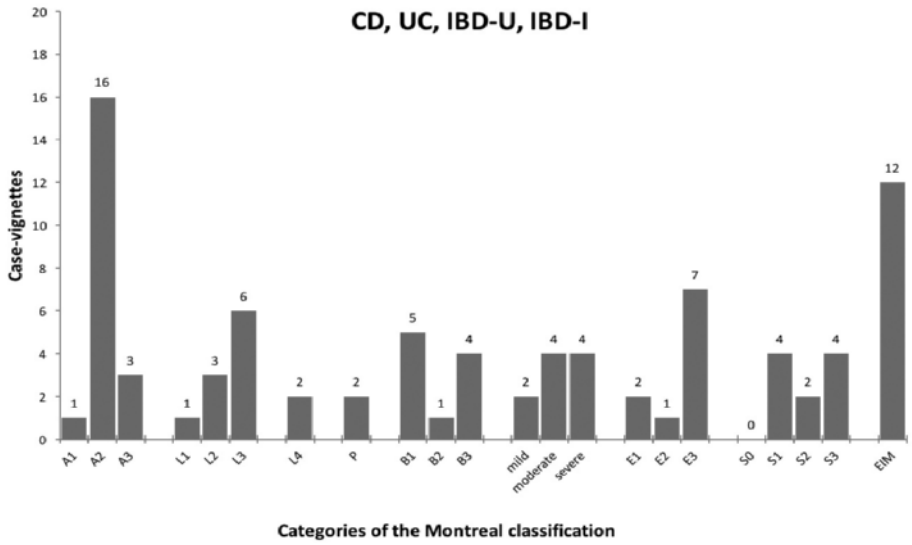


Figure 1 Distribution of the different categories of the Montreal classification for all 20 case vignettes that were scored by 30 observers.
CD: Crohn's disease; IBD: Inflammatory bowel disease; UC: ulcerative colitis; EIM: Extraintestinal manifestations.

The expert panel consisted of two gastroenterologists experienced in IBD care (Dijkstra G, Weersma RK), and one gastroenterologist/PhD in training (Visschedijk MC). The expert panel first assessed the 20 case vignettes separately, discussed their findings and developed an "expert evaluation" for all Montreal items in the 20 case vignettes. This "expert evaluation" was considered as the correct answer. Two additional items were added. Firstly, CD severity was added, because the Montreal classification only allows scoring severity of UC. The Montreal classification does not include any parameters to score severity of CD, therefore observers were asked to give an impression of CD severity (mild, moderate, severe) based on their own clinical experience and judgment. Secondly, observers were asked to score whether any extraintestinal manifestations (EIM) were present.

The 20 case vignettes were sent to 49 observers with different experience levels and professions: gastroenterologists specialized in IBD, gastroenterologists in training and IBD-nurses with a day-to-day experience with IBD patients, all from university and non-university medical centers. The observers received the selected 20 case vignettes, instructions by e-mail and a hyperlink to fill out the online survey (https://www.enquetesmaken.com/), in which the Montreal classification and the two additional items, EIM and CD severity, had to be scored.

The online survey contained the following main items: diagnosis, age of onset and EIM. For the CD case vignettes the observers had to fill in disease location, disease behaviour and disease severity. For the UC case vignettes the observers had to score disease extent and disease severity (Table 1). The diagnosis of CD and UC is standardized and uniformly accepted. However, in 10%-20% of the patients it is difficult to differentiate between CD and UC. These patients are classified as having non-classified chronic colitis (IBD-U). If the pathologist can't differentiate between CD and UC after a colectomy, the patient is classified as having indeterminate colitis (IBD-I).[20-22] Case vignettes with the diagnosis IBD-U or IBD-I are scored as UC, according to the Montreal classification.

Statistical analysis

Statistical analysis was performed using R statistical software. Firstly the inter-observer agreement was calculated using percentages of correct answers. An answer was scored correct if the answer of the observer was identical to the "expert evaluation", percentages of correct answers were calculated for all items.

Secondly Fleiss-kappa (κ) was calculated, which is the standard method to calculate the inter-observer agreement for multiple observers.[23] An observer can only be included in the statistical analysis on the condition that one Montreal item is scored by the observer in all case vignettes. In case of one missing value in one case vignette the observer was excluded from the statistical analysis for this item. The Fleiss-kappa cut-offs were set as follows: < 0.4 poor agreement; 0.41-0.60 moderate agreement; 0.61-0.8 good agreement; > 0.8 excellent agreement.

Subgroup analyses for the inter-observer agreement between profession (gastroenterologist, gastroenterologist in training, IBD-nurse) and level of experience (< 10 years; 10-20 years; > 20 years) were performed by percentages of correct answers. An additional Fisher exact test was used to identify significant differences between the subgroups.

Results

Observers

The online survey was available for six weeks, in which the 49 observers received several reminders. Eventually 30 of the 49 observers completed the survey, a response rate of 61%. Details of the observers are depicted in Table 2. Fifty-four percent of the responders were gastroenterologist and 67% of the observers had less than 10 years experience with IBD patient care.

Table 2 Characteristics of 30 observers *n* (%)

	< 10 yr. of experience with IBD patients	10-20 yr. of experience with IBD patients	> 20 yr. of experience with IBD patients	Non-university center	University medical center	Total
Gastroenterologist	7 (23%)	5 (17%)	4 (13%)	5 (17%)	11 (37%)	16 (54%)
Gastroenterologist in training	10 (33%)			3 (10%)	7 (23%)	10 (33%)
IBD-nurse	3 (10%)	1 (3%)		1 (3%)	3 (10%)	4 (13%)
Total	20 (67%)	6 (20%)	4 (13%)			100%

IBD: Inflammatory bowel disease.

Correct ratings

Average percentage of correctly answered questions for all Montreal and additional items (CD severity and EIM) by different professions was 85%. Age of onset, disease location, perianal disease and disease behaviour in CD had more than 90% correct score over all professions. Disease severity in UC was the worst scored item overall with less than 55% correctly scored by all three professions (Table 3).

Table 3 Percentages of correct answers according to the "expert evaluation" for overall and divided for different professions

	Overall correct answers	Gastro-enterologist	Gastro-enterologist in training	IBD-nurse
Age of onset	94.0%	94.2%	96.8%	86.0%
Diagnosis	96.9%	96.0%	96.0%	98.8%
CD disease Localization	94.0%	90.6%	93.8%	97.5%
CD upper gastrointestinal	91.3%	95.6%	94.0%	84.2%
CD perianal disease	98.0%	99.4%	98.0%	96.7%
CD Disease behaviour	92.4%	92.4%	94.8%	90.0%
CD severity (mild, moderate, severe)	73.9%	68.7%	72.9%	80.0%
UC disease extent	84.0%	89.3%	85.2%	77.5%
UC disease severity colitis	50.7%	53.9%	50.7%	47.5%
EIM	82.1%	82.1%	85.8%	78.5%

EIM: Extraintestinal manifestations; CD: Crohn's disease; UC: ulcerative colitis.

When observers were grouped according to their profession, the additional item severity of the disease (CD), was scored worst by the gastroenterologists (69%) and best by the IBD nurses (80%). IBD-nurses had an excellent correct score on diagnosis (99%) as well as the gastroenterologists (96%) and gastroenterologists in training (96%). According to the fisher exact test, no significant differences were found between the three professions except for scoring of upper gastrointestinal

disease in CD, in which IBD-nurses scored significantly worse than gastroenterologists (P = 0.008) and gastroenterologists in training (P = 0.040).

After calculation of the percentages of correct answers for the observers based on their level of experience, all items of the Montreal classification and the EIM scored above 80%, except for disease severity in UC (48%). The additional item, CD severity, was scored correctly in 70% of cases. Observers with less than 10 years of experience performed best at scoring disease severity (Table 4) and were significantly better at scoring UC severity than observers with 10-20 years (P = 0.003) and more than 20 years (P = 0.003) of experience with IBD patient care. Observers with 10-20 years of experience with IBD patient care were significantly better at scoring upper gastrointestinal disease in CD than observers with less than 10 years (P = 0.007) and more than 20 years (P = 0.007) of experience with IBD patient care.

Table 4 Percentages of correct answers compared to the "expert evaluation" overall and divided for different years of experience with inflammatory bowel disease patients

	Overall correct answers	< 10 yr. of experience	10-20 yr. of experience	> 20 yr. of experience
Age of onset	92.3%	95.2%	93.0%	88.6%
Diagnosis	96.2%	97.0%	93.3%	98.3%
CD disease localization	90.6%	94.2%	90.0%	87.5%
CD upper gastrointestinal	94.8%	91.8%	100.0%	92.5%
CD perianal disease	98.9%	98.5%	98.3%	100.0%
CD disease behaviour	93.0%	92.3%	96.7%	90.0%
CD severity (mild, moderate, severe)	69.9%	73.1%	70.0%	66.7%
UC disease extent	86.6%	86.2%	85.3%	88.3%
UC disease severity colitis	48.0%	56.2%	37.7%	50.0%
EIM	83.4%	82.7%	81.6%	86.0%

EIM: Extraintestinal manifestations; CD: Crohn's disease; UC: ulcerative colitis.

For scoring disease severity in UC, the Montreal requires to score the maximum disease severity ever experienced. Therefore, scoring S0 (meaning clinical remission) would be impossible. We therefore removed observers that scored an S0, and disease severity in UC was calculated again for gastroenterologists and gastroenterologists in training. The percentages of correct answers were 69% and 71%. IBD-nurses were not considered in this analysis because all scored an S0 once or more. Removing S0 for disease severity in UC led to a correct score of 77%, 56% and 56% for observers with less than 10 years, 10-20 years and more than 20 years of experience with IBD patient care.

Inter-observer agreement

The inter-observer agreement was excellent for diagnosis (κ = 0.96), CD location (κ = 0.82) and perianal disease (κ = 0.91). Age of onset (κ = 0.67) and upper gastrointestinal disease (κ = 0.62) were scored with a good inter-observer agreement. Disease severity was scored poorly (κ = 0.23) for UC. The additional clinical item, CD severity, was scored with moderate concordance (κ = 0.44). In total there were 19 EIMs in 12 case vignettes. The inter-observer agreement for occurrence of EIM was good (κ = 0.68) (Table 5).

By removing all the observers that stated an S0 once or more, only 7 observers remained which led to a kappa of 0.57, resulting in moderate inter-observer agreement for severity in UC. Kappa was also calculated again for disease extent and disease severity in UC, but now for 30 observers with all the missing values being replaced by the correct answer as established by the "expert evaluation". No significant differences in the inter-observer agreement for 30 and 20/21 observers scoring disease severity and disease extent in UC were found.

Table 5 Inter-rater agreement kappa for all items of all categories of Montreal classification

Item Montreal classification	Overall kappa
Age of onset	0.67 (*n* = 28)
Diagnosis	0.96 (*n* = 28)
CD disease localization	0.82 (*n* = 28)
CD upper gastrointestinal	0.62 (*n* = 28)
CD perianal disease	0.91 (*n* = 28)
CD disease behaviour	0.79 (*n* = 26)
CD severity (mild, moderate, severe)	0.44 (*n* = 24)
UC disease extent	0.65 (*n* = 21)
UC disease severity colitis	0.23 (*n* = 20)
EIM	0.68 (*n* = 28)

Observers were only included if they scored at least one item in all case-vignettes.
EIM: Extraintestinal manifestations; CD: Crohn's disease; UC: ulcerative colitis.

Discussion

The aim of this study was to assess the validity of accurate phenotyping using the Montreal IBD classification system with 2 additional items (CD severity and EIM) for both CD and UC within the Netherlands.

According to our study, the Montreal is a reliable classification system for phenotypes in IBD, except for disease severity in UC. The assessment of disease severity for UC as described in the Montreal classification system is difficult in the case of retrospective chart reviews. Since severity

in CD is not a classification item in the Montreal, we asked the observers to score CD severity based on their personal interpretation of the case vignettes. This resulted in a low consistency between observers, but this item was scored with higher concordance (with fewer instructions) than disease severity in UC.

Until now only limited data on the reliability and reproducibility of the Montreal classification is available. An Australian-New Zealand and United States study[18,19] found a good inter-observer agreement for CD, however the Australian-New Zealand study did not include the scoring of UC and neither study included an assessment of disease severity for both UC and CD. In our study the inter-observer agreement for diagnosis was excellent (κ = 0.96), which was comparable to the Australian-New Zealand study (κ = 0.82). The inter-observer agreement for age of onset was only "good" in our cohort (κ = 0.67) as compared to excellent in the Australian-New Zealand (κ = 0.84) and the US study (κ = 0.98). The observers in our cohort were better at scoring disease localization in CD, upper gastrointestinal involvement, perianal disease and disease behaviour in CD. Disease extent in UC was similarly scored in our cohort (κ = 0.65) as in the Australian-New Zealand study (κ = 0.67).[18,19]

Classifying disease severity in patients' records ("real life") is still a problem because of missing or unclear descriptions. Clinicians should strive to be complete and accurate in their medical reporting. A clearer definition of disease severity is needed because apparently there is no consensus between clinicians about mild, moderate or severe disease in (real life) patients. For disease severity there are several classification systems e.g., CD activity index[24] and the UC activity index[25] that assess disease severity by clinical symptoms, however these symptoms are present at a specific time point and cannot be assessed in a retrospective manner. The CD digestive damage score (Lemann score) is a measurement for cumulative structural bowel damage, assessed by scoring disease severity for damage location, severity, extent, progression and reversibility, diagnosed by image modalities and the history of surgical resections. Ultimately a prediction model gives a reflection of progressive and destructive disease course.[26] The Lemann score might be a good instrument for classifying disease severity.

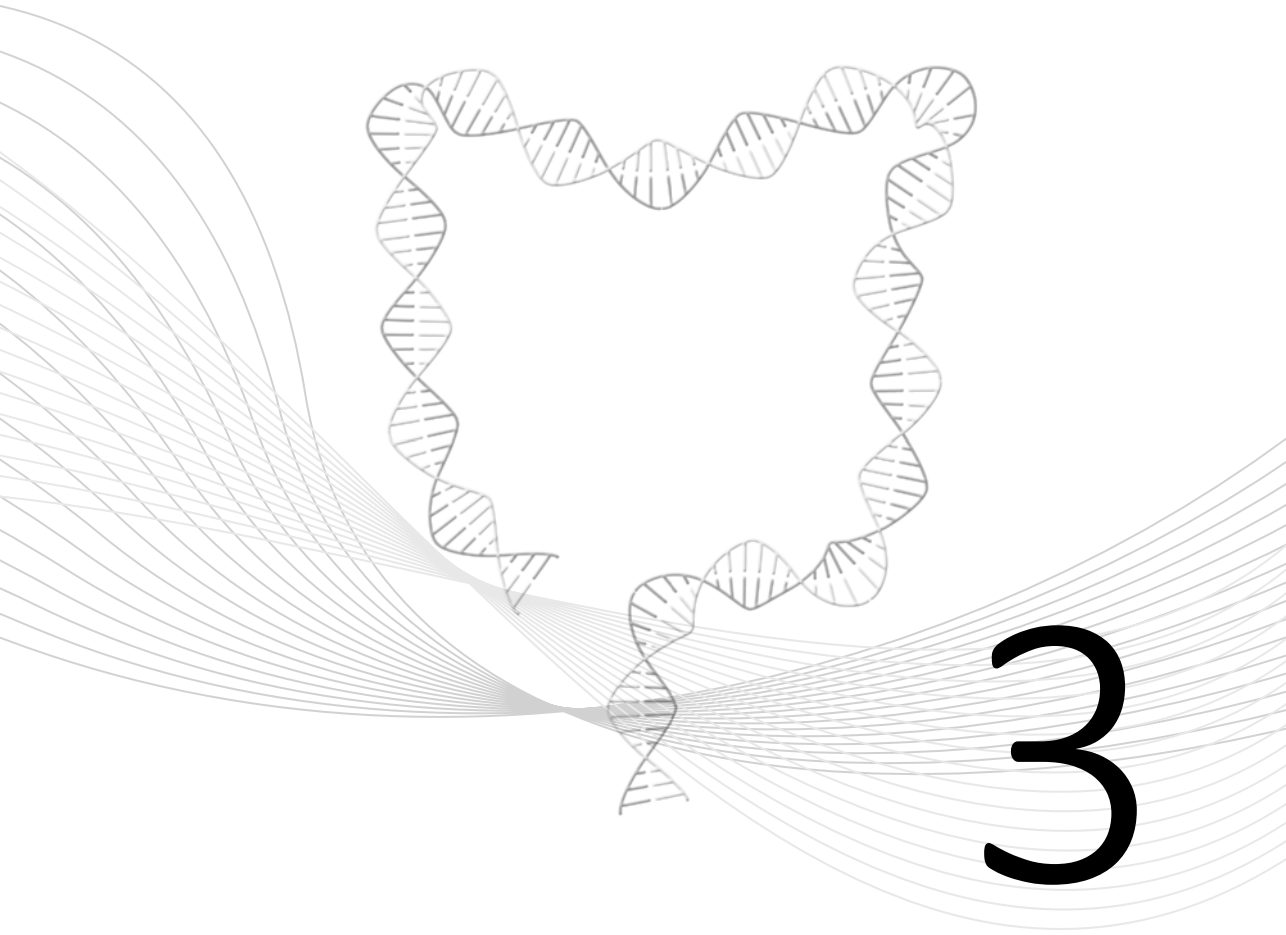
Since IBD is a chronic disease with unpredictable disease behaviour, it is very important that clinicians can identify those individuals with a severe disease course, risk of side effects to therapy or those who would benefit from lifestyle or environmental changes. It is expected that molecular and/or pharmaco-genetic markers will play an increasing role in predicting disease course or response to medication in the future.[27] A good opportunity to predict individual disease behaviour is by linking their uniform phenotypic characteristics with our knowledge of the molecular basis of the disease. In IBD research an increasing number of biobanks are being set up worldwide allowing for linking molecular data to phenotypic data. To ensure high-quality data, validation of the Montreal classification is mandatory for these kinds of multicenter prospective data collections.

This Dutch validation study has a larger observer group than the previously mentioned studies. It is the first to include UC and CD disease severity, and to differentiate between professions. We found a good inter-observer agreement for diagnosis, localization, disease behaviour, disease extent and the occurrence of EIM. The reliability for assessment of disease severity for UC was poor, and moderate for the additional CD severity item. Optimal reporting of uniform phenotypes of patient cohorts is of utmost importance, especially in genetic and clinical research. Uniform phenotyping will ultimately allow for integration of clinical phenotypes with high-throughput–omics data (integration of genetic, expression or metagenomic data), which will increase our understanding of IBD pathogenesis, and allow for better patient stratification and classification.

References

1. Loftus EV. Clinical epidemiology of inflammatory bowel disease: Incidence, prevalence, and environmental influences. *Gastroenterology* 2004;126:1504-1517.
2. Molodecky NA, Soon IS, Rabi DM, *et al*. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology* 2012;142:46-54.e42; quiz e30.
3. Bernstein CN, Fried M, Krabshuis JH, *et al*. World Gastroenterology Organization Practice Guidelines for the diagnosis and management of IBD in 2010. *Inflamm Bowel Dis* 2010;16:112-124.
4. Sands BE. From symptom to diagnosis: clinical distinctions among various forms of intestinal inflammation. *Gastroenterology* 2004;126:1518-1532.
5. Lakatos PL, Lakatos L, Kiss LS, *et al*. Treatment of extraintestinal manifestations in inflammatory bowel disease. *Digestion* 2012;86 Suppl 1:28-35.
6. Mowat C, Cole A, Windsor A, *et al*. Guidelines for the management of inflammatory bowel disease in adults. *Gut* 2011;60:571-607.
7. D'Haens GR, Panaccione R, Higgins PD, *et al*. The London Position Statement of the World Congress of Gastroenterology on Biological Therapy for IBD with the European Crohn's and Colitis Organization: when to start, when to stop, which drug to choose, and how to predict response? *Am J Gastroenterol* 2011;106:199-212; quiz 213.
8. Van Assche G, Dignass A, Panes J, *et al*. The second European evidence-based Consensus on the diagnosis and management of Crohn's disease: Definitions and diagnosis. *J Crohns Colitis* 2010;4:7-27.
9. Dignass A, Eliakim R, Magro F, *et al*. Second European evidence-based consensus on the diagnosis and management of ulcerative colitis part 1: definitions and diagnosis. *J Crohns Colitis* 2012;6:965-990.
10. Abraham C, Cho JH. Inflammatory bowel disease. *N Engl J Med* 2009;361:2066-2078.
11. Xavier RJ, Podolsky DK. Unravelling the pathogenesis of inflammatory bowel disease. *Nature* 2007;448:427-434.
12. Jostins L, Ripke S, Weersma RK, *et al*. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;491:119-124.
13. Khor B, Gardet A, Xavier RJ. Genetics and pathogenesis of inflammatory bowel disease. *Nature* 2011;474:307-317.
14. Anderson CA, Boucher G, Lees CW, *et al*. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet* 2011;43:246-252.
15. Franke A, McGovern DP, Barrett JC, *et al*. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet* 2010;42:1118-1125.
16. Gasche C, Scholmerich J, Brynskov J, *et al*. A simple classification of Crohn's disease: report of the Working Party for the World Congresses of Gastroenterology, Vienna 1998. *Inflamm Bowel Dis* 2000;6:8-15.
17. Silverberg MS, Satsangi J, Ahmad T, *et al*. Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. *Can J Gastroenterol* 2005;19 Suppl A:5A-36A.
18. Dassopoulos T, Nguyen GC, Bitton A, *et al*. Assessment of reliability and validity of IBD phenotyping within the National Institutes of Diabetes and Digestive and Kidney Diseases (NIDDK) IBD Genetics Consortium (IBDGC). *Inflamm Bowel Dis* 2007;13:975-983.
19. Krishnaprasad K, Andrews JM, Lawrance IC, *et al*. Inter-observer agreement for Crohn's disease sub-phenotypes using the Montreal Classification: How good are we? A multi-centre Australasian study. *J Crohns Colitis* 2012;6:287-293.
20. Price AB. Overlap in the spectrum of non-specific inflammatory bowel disease--'colitis indeterminate'. *J Clin Pathol* 1978;31:567-577.
21. Wells AD, McMillan I, Price AB, *et al*. Natural history of indeterminate colitis. *Br J Surg* 1991;78:179-181.
22. Zhou N, Chen WX, Chen SH, *et al*. Inflammatory bowel disease unclassified. *J Zhejiang Univ Sci B* 2011;12:280-286.

23. Fleiss JL. Measuring nominal scale agreement among many raters. *Psychol Bull* 1971;76:378-382.
24. Best WR, Becktel JM, Singleton JW, *et al*. Development of a Crohn's disease activity index. National Cooperative Crohn's Disease Study. *Gastroenterology* 1976;70:439-444.
25. Sutherland LR, Martin F, Greer S, *et al*. 5-Aminosalicylic acid enema in the treatment of distal ulcerative colitis, proctosigmoiditis, and proctitis. *Gastroenterology* 1987;92:1894-1898.
26. Pariente B, Cosnes J, Danese S, *et al*. Development of the Crohn's disease digestive damage score, the Lémann score. *Inflamm Bowel Dis* 2011;17:1415-1422.
27. Festen EA, Weersma RK. How will insights from genetics translate to clinical practice in inflammatory bowel disease? *Best Pract Res Clin Gastroenterol* 2014;28:387-397.



3

Cohort profile: design and first results of the Dutch IBD Biobank: a prospective, nationwide biobank of patients with inflammatory bowel disease

LM Spekhorst*, F Imhann*, EAM Festen, AA van Bodegraven, NKH de Boer, G Bouma, HH Fidder, G D'Haens, F Hoentjen, DW Hommes, DJ de Jong, M Löwenberg, PWJ Maljaars, AE van der Meulen-de Jong, B Oldenburg, MJ Pierik, CY Ponsioen, PC Stokkers, HW Verspaget, MC Visschedijk, CJ van der Woude, G Dijkstra, RK Weersma; on behalf of the Parelsnoer Institute (PSI) and the Dutch Initiative on Crohn and Colitis (ICC)

*Authors contributed equally

BMJ Open 2017 Nov 8;7(11)

Abstract

Purpose: The Dutch IBD Biobank aims to facilitate the discovery of predictors for individual disease course and treatment response in patients with inflammatory bowel disease (IBD). In this paper, we aim to describe the establishment of the Dutch IBD Biobank, including the facilitators and barriers to establishment. Moreover, we aim to provide a complete overview of the content of the Dutch IBD Biobank.

Participants: Since 2007, every patient with IBD treated in one of the eight Dutch university medical centres is asked to participate in the Dutch IBD Biobank in which 225 standardised IBD-related data items and biomaterials, such as serum, DNA, biopsies and a stool sample, are collected.

Findings to date: As of June 2014, the Dutch IBD Biobank had enrolled 3388 patients with IBD: 2118 Crohn's disease (62.5%), 1190 ulcerative colitis (35.1%), 74 IBD- unclassified (2.2%) and 6 IBD-indeterminate (0.2%). The inclusion of patients with IBD is ongoing. The quality of the biomaterials is good and serum, DNA and biopsies have been used in newly published studies.

Future plans: The genotyping (750,000 genetic variants) of all participants of the Dutch IBD Biobank is currently ongoing, enabling more genetic research. In addition, all participants will start reporting disease activity and outcome measures using an online platform and mobile app.

Introduction

Inflammatory bowel disease (IBD) is a chronic inflammatory disease of the gut comprising Crohn's disease (CD) and ulcerative colitis (UC). Of the 17 million inhabitants in the Netherlands, 39,000 individuals have been diagnosed with CD and 48,000 individuals with UC.[1] Approximately 39 new individuals per 100,000 are newly diagnosed with IBD every year. This incidence rate continues to rise, posing an increasing burden on society.[2] The clinical symptoms of IBD consist of diarrhoea, abdominal discomfort, weight loss, fatigue and rectal bleeding. However, these symptoms vary greatly both between individuals and in time. Some patients with IBD have a relatively mild disease course, requiring only limited therapeutic intervention, while others have a severe disease course with frequent flares requiring expensive medical and surgical interventions.

In recent years, many case–control studies have been performed to identify factors that can explain the *onset* of IBD. Genome-wide association studies (GWAS) have identified 200 genomic loci that are involved in the onset of IBD.[3] Epidemiological studies have identified environmental risk factors including smoking, appendectomy, infections, antibiotics, diet and lifestyle (stress, lack of sleep and/or exercise) that could trigger the onset of IBD.[4] Studies on the bacterial composition of the gut (the gut microbiota) have identified distinct microbial compositions associated with IBD.[5,6] Unfortunately, these studies provide little insight into reasons for the heterogeneous clinical presentation and *disease course* of patients with IBD. As a consequence, limited progress has been made in translating basic science into personalised treatment. Predicting individual disease outcome and tailoring IBD treatment requires prospective patient data on disease activity, complications and treatment, as well as biomaterials and *-omics* data (genome, transcriptome and gut microbiome), in order to link biomarkers to disease. To this aim, the prospective Dutch IBD Biobank was created. A new national institute to facilitate the biobank and other national biobanks was founded by the Dutch Federation of University Medical Centres (NFU) in 2007 and called the Parelsnoer Institute (PSI).[7] Gastroenterologists who specialised in treating patients with IBD in all eight Dutch university medical centres (UMC), together with a team of information architects and laboratory experts, built up the Dutch IBD Biobank.

The main objective of the biobank is to facilitate the discovery of predictors (both epidemiological risk factors and biomarkers) for individual disease course and treatment response, by:

1. providing full clinical records of patients describing their individual disease course over a prolonged period of time;
2. providing high-quality biomaterials;
3. standardising patient data collection and questionnaires during outpatient clinic visits and thereby improving clinical care.

The aim of this paper is to inform the IBD research community about the existence of the Dutch IBD Biobank and to give an elaborate overview of the establishment process as well as the content.

Cohort description

Design, participating centres and the Dutch healthcare setting

The Dutch IBD Biobank is a prospective, nationwide biobank in which both data and biomaterials are collected. In the Netherlands, there are approximately 80 hospitals and 8 UMCs (tertiary referral centres), where patients with complex IBD are referred to. All eight Dutch UMCs participate in the Dutch IBD Biobank. The Dutch UMCs are: the Amsterdam Medical Centre in Amsterdam, the Erasmus Medical Centre in Rotterdam, the Leiden University Medical Centre in Leiden (LUMC), the Maastricht University Medical Centre in Maastricht (MUMC), the Radboud University Nijmegen Medical Centre in Nijmegen, the University Medical Center Groningen in Groningen (UMCG), the University Medical Centre Utrecht in Utrecht and the VU (*Vrije Universiteit*) University Medical Centre in Amsterdam. PSI and the Dutch IBD Biobank are part of the Biobanking and Biomolecular Resources Research Infrastructure of the Netherlands (BBMRI-NL). This is the Dutch national node of BBMRI-ERIC, the largest research infrastructure project in Europe.[8]

Standardised data collection: the information model

Gastroenterologists from each of the eight UMCs convened to design the information model based on literature review and clinical standards. A working group of gastroenterologists made a longlist of data items including a definition for each data item. This longlist was subsequently discussed during a meeting in 2006, where one or more representatives from each Dutch UMC were present. Data items and definitions were accepted, modified if deemed necessary, or rejected if deemed not part of the core data set. This process was repeated until consensus was reached. The Dutch IBD Biobank prospectively collects 225 standardised data items on various topics, including patient demographics, family history, diagnosis, disease activity, disease localisation, results of physical examinations, radiographic imaging results, laboratory and endoscopy results, previous and current treatment, as well as a wide array of disease and treatment complications. Validated questionnaires and scores, such as the Harvey-Bradshaw Index (HBI), the Simple Clinical Colitis Activity Index (SCCAI) and the Montreal classification are incorporated in the information model. This model contains both the IBD-related items as well as instructions on how to score these items. It has been shown that clinicians score subphenotypes of IBD similarly, with a good to excellent interobserver agreement.[9] The information model is provided in English in online **Supplementary Table 1** and can be downloaded in Dutch on the PSI website: www.parelsnoer.org. The Dutch IBD

Biobank information model is regularly updated. The latest version is based on the coding system called Detailed Clinical Models (http://www.detailedclinicalmodels.nl/dcm-en) and is called PRISMA (Parelsnoer Repository for Information Specification, Modelling, and Architecture).

Local databases and infrastructure

Each UMC has implemented the information model and collects and stores their patient information locally. As stated by the NFU, research data should be collected and registered directly at the source, that is, during the patient visit. Therefore, the data collection process should be incorporated into the clinical care structure.[10] This approach has been gradually implemented in each UMC depending on the capacities of their electronic health record (EHR) system. At the moment, each UMC has a procedure to extract, transform and upload pseudonymised information of participants to the PSI central database (Figure 1). The UMCs are in different stages of having implemented the 'at the source' approach. In some UMCs it is already fully implemented, whereas in other UMCs this process has not yet begun. The first visit is prepared by a trained research nurse and since most of the 225 data items do not change during every visit, for example, family history, medical doctors usually only need to register a subset of items during visits.

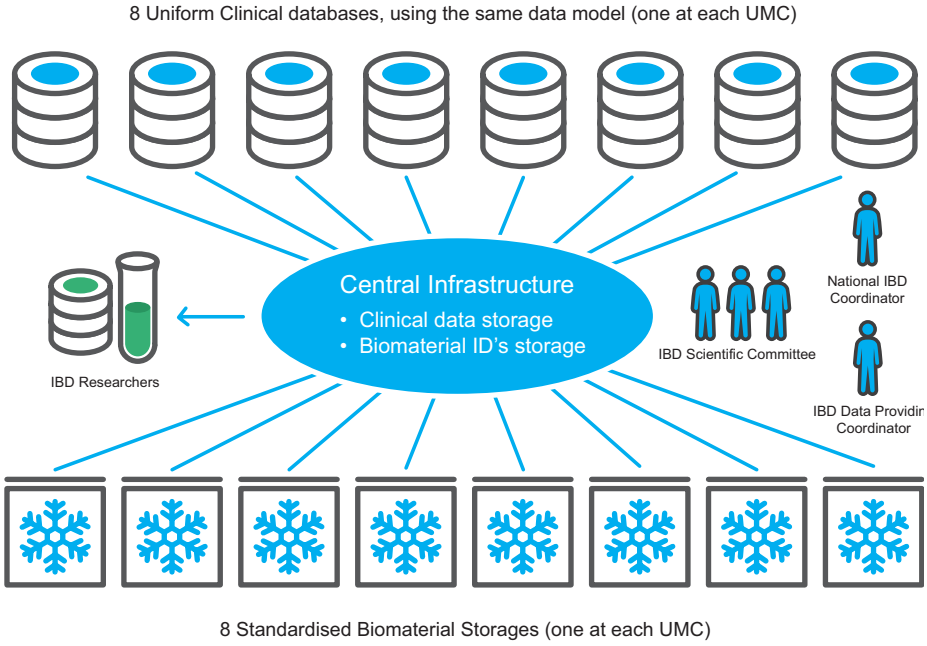


Figure 1 Overview of the data and biomaterial infrastructure of the Dutch IBD Biobank, built by the Parelsnoer Institute in collaboration with all eight university medical centres (UMC) in the Netherlands.
IBD: inflammatory bowel disease; ID: identifier.

Central Database and Central Data Infrastructure

Pseudonymised information about study participants is stored in the Central Database, managed by the Advanced Data Management (ADM) section of the Department of Medical Statistics and BioInformatics of the LUMC.

The software ProMISe, a web-based relational database management system for the design, maintenance and use of clinical data management, is used to store the Central Database. (https://www.msbi.nl/promise/). Researchers can access data in the Central Database following approval of their research proposal in secure web-based environment. Together, the Central Database and the web application form the Central Data Infrastructure (Figure 1).[7]

Data upload and pseudonymisation

In each UMC, data are automatically uploaded from the Local Database to the Central Database at least once a month. During the upload process, pseudoanymisation is performed by a trusted third party (TTP). Only the TTP has access to key containing both the local identifiers and the Dutch IBD Biobank identifier. Prior to the upload, data validation is performed locally on a set of essential data items. If necessary, corrections are made locally and subsequently included in the next upload. A full audit trail is in place for the entire process.

Privacy and information security audits

ADM, the Central Database and the Central Data Infrastructure software are audited according to Dutch NEN7510[11] international ISO 27.001[12] information security guidelines. ADM is audited twice per year while its software is periodically audited by Lloyds Register Quality Assurance, a certified independent auditor.

Biomaterial collection

In addition to the data items, biomaterials are collected from all patients with IBD: including DNA, serum, faeces, mucosal biopsies and resection specimens when surgical procedures were required. Laboratory experts of all eight university hospitals convened to create uniform biomaterial collection and processing protocols. The biomaterials are stored in one of the eight local biobanks (Figure 1). The biomaterial identifiers are uploaded to the Central Database and linked to the clinical data. Neither the local biomaterial identifiers nor the stickers on the biomaterial vials contain identifiable patient information. During the upload process, a unique additional biomaterial identifier is added to the local biomaterial identifier in case multiple UMCs have a biomaterial with the same identifier. When a research project is approved, all eight local biobanks will send the required biomaterials to the researcher while the biomaterial identifiers linked to the clinical data can be downloaded using the secure web portal of the Central Infrastructure. If a biomaterial sample does not meet

the required standards, the sample will be disposed. A brief summary of the biomaterial protocol is provided in Table 1.[7] The entire biomaterial protocols can be downloaded from www.parelsnoer.org, but are only available in Dutch.

Coordination

The Dutch IBD Biobank has two national coordinators and an assistant coordinator, who manage updates of the information model and the delivery of data and biomaterial to researchers (Figure 1).

Table 1 Sample collection[7]

Sample	Volume/ number	Processing	Time	Aliquoting	Storage	Additional information
Serum	10 ml clotted blood	2000xg at room temperature or 4°C for 10 minutes	Within 2-4 hours	≥ 5 x 0.5 ml	–80°C	Deviations
DNA	10 ml EDTA blood	Cell pellet, to UMC specifications	Within 4 weeks (4°C) or 3 months (< -20°C)	≥ 2 stock aliquots	4°C or lower	OD-ratio 260/280 and concentration in µg/ml
Faeces	Not defined	Direct storage or after homogenization	Within 12 hours	≥ 5 x 5 gr	–80°C	None
Intestinal biopsy	2 per localization: 'normal' and 'affected/ inflamed'	Formalin fixation and paraffin embedding	Immediate	Per set	Room temperature	None
Resection specimen	2 per localization: 'normal' and 'affected/ inflamed'	Formalin fixation and paraffin embedding	At Pathology	0.5 cm^3 samples	Room temperature	Only if feasible
Resection specimen	2 per localization: 'normal' and 'affected/ inflamed'	Snap frozen in isopentane	At Pathology	0.5 cm^3 samples	–80°C	Only if feasible

OD: optical density; UMC: university medical centre.

Informed consent

All patients with IBD who are treated in the Dutch UMCs are asked to participate in the Dutch IBD Biobank by their gastroenterologist during a visit to the outpatient department of their UMC. If they are willing to participate, they are asked to sign an informed consent form (English translation in

online Supplementary Document 1). Patients who choose to participate may revoke their consent at any point, after which their data and biomaterials will be removed from the Dutch IBD Biobank. Data and biomaterials that have already been sent to a researcher cannot be revoked, which is clearly stated in the patient informed consent form.

Patient enrolment

Patient enrolment started in January 2007 and is ongoing (Table 2).

Not all patients were asked to join at once, but they were asked in batches so gastroenterologist and research nurses could manage the initial data registration.

Every patient with IBD enrolled has a proven IBD diagnosis according to the Lennard-Jones criteria.[13] Diagnosis is confirmed by endoscopy, radiology and/or histology.

Table 2 Demographic characteristics of patients with IBD after the first data download on 17 July 2014, per university medical centre

	Total	MUMC	VUMC	AMC	UMCG	UMCU	EMC	LUMC	UMCN
n	3388	373	369	405	625	524	260	458	374
CD	2118	219	206	264	344	337	194	310	244
UC/IBD-U /IBD-I	1270	154	163	141	281	187	66	148	130
Sex (f/m%)	59/41	54/46	64/36	57/43	59/41	58/42	64/36	58/42	64/36
Age at diagnosis*	26 (20-37)	31 (22-44)	28 (21-37)	26 (20-35)	27 (21-39)	25 (19-35)	23 (18-30)	26 (20-34)	27 (20-37)
Disease duration*	12 (5-20)	8 (2-17)	11 (6-20)	13 (6-22)	8 (4-15)	14 (6-24)	12 (6-20)	15 (7-23)	14 (7-24)

*Median years with 25%–75% IQR.

AMC: Amsterdam Medical Centre; CD: Crohn's disease; EMC: Erasmus Medical Centre; F: female; IBD: inflammatory bowel disease; IBD-I: inflammatory bowel disease-indeterminate; IBD-U: inflammatory bowel disease-unclassified; LUMC: Leiden University Medical Centre; M: male; MUMC: Maastricht University Medical Centre; UC: ulcerative colitis; UMCG: University Medical Center Groningen; UMCN: Radboud University Nijmegen Medical Centre; UMCU: University Medical Centre Utrecht; VUMC: VU (Vrije Universiteit) University Medical Centre (Amsterdam).

Definitions

To create an overview of the content of the biobank, the characteristics of the patients were assessed. The following clinical and demographic items reported in this study are registered at the time of inclusion in the Dutch IBD Biobank and are referred to as *baseline*: first diagnosis, disease localisation, smoking status, employment status, gender, ethnicity, presence of a stoma or pouch, disease activity (modified HBI and modified SCCAI score) and date of birth. Disease localisation is scored according to the Montreal classification, which describes the maximum disease extent during entire disease course, and is registered at *baseline*. Disease localisation has to be confirmed by radiology, endoscopy or histology assessment. The items dysplasia, bowel

cancer, family history of IBD, current diagnosis and medication use described in this study were registered *during the last follow-up visit before the data download* in July 2014. Items describing disease behaviour, surgery, appendectomy, extraintestinal manifestations (EIM) and complications were registered *over the entire disease course up to baseline*. The definitions *baseline*, *last follow-up visit before the data download* and *over the entire disease course up to baseline* are graphically explained in online **Supplementary Figure 1**.

Statistical analyses

All descriptive statistics and statistical analyses are performed using Stata software V.13.1 (http://www.stata.com/). Continuous variables are expressed as medians and IQRs 25 and 75. Qualitative variables are presented as counts and frequencies. We compared outcomes between patients with CD and UC. Qualitative variables were analysed using the Pearson's X^2 test. Quantitative variables were analysed using the Mann-Whitney U test. We performed a multivariate analysis of the effect of smoking on different outcomes in all patients with IBD. We corrected for covariates with $P < 0.20$ in the univariate analyses (age, gender, diagnosis, disease duration and prior anti-tumour necrosis factor use). The statistical models were built using backward selection: covariates that were not statistically significantly influencing the outcome variable ($P > 0.05$) were removed from the model. We then applied the same strategy to patients with CD and UC separately to correct for disease activity. A P value < 0.05 was considered statistically significant.

Follow-up

Clinical and demographical follow-up data are collected at every visit to an outpatient department. Usually, patients with IBD in the Netherlands are seen by a gastroenterologist twice a year. This is standard clinical care following treatment protocols used in every UMC. The disease course is heterogeneous, as a consequence, data available on follow-up can be extensive for one patient but more limited for another. If requested by the gastroenterologist, a blood sample is taken. Furthermore, if required, intestinal mucosal biopsies are collected during endoscopy and resection specimens are obtained during surgery.

Findings to date

Consent rate and differences between participants and non-participants

We first assessed possible differences between patients with IBD willing to participate in the Dutch IBD Biobank and patients with IBD who did not want to participate. To do so, a subset at one UMC (UMCG) was downloaded and analysed. This subset was used because privacy guidelines do not

allow data of participants not wishing to take part to be uploaded to the PSI central database. On 17 July 2014, after the first data download, 786 patients were asked to participate in the UMCG. Of these, 742 patients with IBD gave their informed consent while 44 patients with IBD declined to participate. The consent rate was 93.4%. Table 3 provides an overview of the characteristics of those who consented to participate and those who did not. Of the 742 patients who consented, 625 were used in the analysis of the 2014 data because they met the selection criteria (clear IBD diagnosis, known date of birth and gender, informed consent and isolated DNA available including a biomaterial identifier). The characteristics of the consenting and non-consenting patients were similar. Only disease location according to the Montreal classification was statistically significantly different between these two groups (P = 0.037, X^2 test).

Table 3 Baseline characteristics of the responders and non-responders recruited through the University Medical Center Groningen on 17 July 2014

Responders			
	IBD (CD, UC, IBD-U)	CD	UC
n (%)	742 (100%)	411 (55%)	294 (40%)
Sex	742 (100%)	411 (100%)	294 (100%)
Male	305 (41%)	141 (34%)	142 (48%)
Female	437 (59%)	270 (66%)	152 (52%)
Age of onset median years (IQR 25-75)	26.8 (20-38)	24.5 (19-35)	30.6 (23-41)
Disease duration at inclusion median years (IQR 25-75)	8.2 (4-15)	9.3 (4-15)	7.6 (4-14)
Disease location (according Montreal)			
Crohn's disease		411 (100%)	
A1: diagnosis ≤ 16 years		58 (14%)	
A2: diagnosis 17-40 years		278 (68%)	
A3: diagnosis > 40 years		75 (18%)	
L1: ileal disease[a]		148 (37%)	
L2: colonic disease[a]		85 (22%)	
L3: ileocolonic disease[a]		163 (41%)	
L4: upper GI disease[b]		41 (10%)	
P: perianal		130 (32%)	
B1: non-stricturing, non-penetrating		211 (51%)	
B2: stricturing		134 (33%)	
B3: penetrating		66 (16%)	
Ulcerative colitis			288 (100%)
E1: proctitis			40 (14%)
E2: left-sided colitis			92 (32%)
E3: extensive colitis			156 (54%)

Table 3 *continued*

Non-responders			
	IBD (CD, UC, IBD-U)	CD	UC
n (%)	44 (100%)	25 (57%)	16 (36%)
Sex	44 (100%)	25 (100%)	16 (100%)
Male	16 (36%)	9 (36%)	5 (31%)
Female	28 (64%)	16 (64%)	11 (69%)
Age of onset median years (IQR 25-75%)	30.3 (19-42)	19.6 (17-39)	33.3 (25-42)
Disease duration at inclusion median years (IQR 25-75%)	8.1 (4-12)	7.2 (3-12)	8.8 (5-13)
Disease location (according to Montreal guidelines)			
Crohn's disease		25 (100%)	
A1: diagnosis ≤ 16 years		7 (28%)	
A2: diagnosis 17-40 years		12 (48%)	
A3: diagnosis > 40 years		6 (24%)	
L1: ileal disease*		4 (16%)	
L2: colonic disease*		10 (40%)	
L3: ileocolonic disease*		11 (44%)	
L4: upper gastrointestinal disease		0 (0%)	
P: perianal		9 (36%)	
B1: non-stricturing, non-penetrating		11 (44%)	
B2: stricturing		10 (40%)	
B3: penetrating		4 (16%)	
Ulcerative colitis			15 (100%)
E1: proctitis			5 (33%)
E2: left-sided colitis			5 (33%)
E3: extensive colitis			5 (33%)

[a]These percentages were calculated for 396 patients with CD (responders);
[b]These percentages were calculated for 402 patients with CD (responders);
*P = 0.037.
CD: Crohn's disease; GI: gastrointestinal; IBD: inflammatory bowel disease; IB-U: inflammatory bowel disease-unclassified; UC: ulcerative colitis.

The characteristics of the Dutch patients with IBD in UMCs

A download of data on 17 July 2014 was analysed to explore the demographic and clinical characteristics of the cohort recruited to that date. It included 3388 patients with IBD: 2118 CD (62.5%), 1190 UC (35.1%), 74 IBD-unclassified (2.2%) and 6 IBD-indeterminate (0.2%). The median age of patients with IBD at inclusion was 42 years old (IQR 32–54 years) (Tables 4-6). In all, 93% of patients are of Central European Caucasian descent and the other 7% are of African, Hindustani, Moroccan, Turkish, Asian, Jewish, other western, other non-western or mixed descent. Smoking status at the time of first IBD diagnosis was registered for 3021 patients with IBD (89%), and more patients with CD smoked compared with patients with UC (44% CD, 18% UC, P < 0.001). Patients with UC were more likely to have quit smoking in the 6 months prior to the first IBD diagnosis (1.0% CD,

4% UC, P < 0.001). Ileocolonic disease in patients with CD (46%) (Figure 2) and extensive colitis (E3) in patients with UC (56%) (Figure 3) are more common in our cohort than in other studies (Figures 4 and 5).[14-18] The high number of patients with extensive disease in our cohort can be explained by a selection bias (tertiary referral centres). The disease locations in CD were similar in men and women (Figure 6).

Table 4 Demographic characteristics of patients with inflammatory bowel disease in the Dutch IBD Biobank cohort on 17 July 2014

	IBD (CD, UC, IBD-I, IBD-U)	CD	UC
n (%)	3388 (100%)	2118 (62%)	1190 (35%)
Sex	3388 (100%)	2118 (100%)	1189 (100%)
Male	1377 (41%)	773 (36%)*	566 (48%)*
Female	2010 (59%)	1345 (64%)*	623 (52%)*
Age at inclusion median years (IQR 25-75)	42.5 (32-54)	41.1 (31-53)*	45.5 (34-56)*
Ethnicity	3323 (100%)	2073 (100%)	1170 (100%)
Caucasian	3090 (93%)	1930 (93%)	1084 (93%)
Other	233 (7%)	143 (7%)	86 (7%)
Non-IBD surgery			
Appendectomy†	394 (12%)	313 (15%)*	76 (6%)*
Smoking status at diagnosis	3021 (100%)	1910 (100%)	1037 (100%)
Current smoker	1052 (35%)	846 (44%)*	190 (18%)*
Former smoker (< 6 mth)	60 (2%)	19 (1.0%)*	40 (4%)*
Former smoker (> 6 mth)	601 (20%)	254 (13%)*	328 (32%)*
Never smoked	1308 (43%)	791 (42%)*	479 (46%)*

†Missing values were scored as absent;
*P < 0.001.
CD: Crohn's disease; GI: gastrointestinal; IBD: inflammatory bowel disease; IBD-I: inflammatory bowel disease-indeterminate; IBD-U: inflammatory bowel disease-unclassified; UC: ulcerative colitis.

Moreover, the most extensive disease during the entire disease duration (Montreal L (disease location) in patients with CD and Montreal E (disease extent) in patients with UC) is well documented in the Dutch IBD Biobank, while other studies often only report disease extent at the time of diagnosis (median disease duration in the Dutch IBD Biobank is 12 years). EIMs are more common in patients with CD than in patients with UC, which we corroborated in the Dutch IBD Biobank data (Figure 7).[19-21] We found that patients with UC who smoked more often suffered from ocular manifestations and arthropathy than those who did not smoke, matching previous findings.[22,23] An increased risk of EIM in patients with CD who smoked has previously been reported,[24] but we could not confirm this result in our cohort.

Table 5 Clinical characteristics, extraintestinal manifestations and complications in patients with inflammatory bowel disease in the Parelsnoer Institute cohort

	IBD	CD	UC
n (%)	3388 (100%)	2118 (62%)	1190 (35%)
Disease Characteristics			
Age of onset median years (IQR 25-75)	26.4 (20-37)	24.6 (19-33)**	30.1 (22-41)**
Disease duration at inclusion median years (IQR 25-75)	11.5 (5-20)	12.2 (6-22)**	10.7 (5-19)**
Family history of IBD	932 (28%)	613 (29%)*	301 (25%)*
Disease location (Montreal classification)			
L1: ileal disease[a]		379 (23%)	
L2: colonic disease[a]		518 (31%)	
L3: ileocolonic disease[a]		780 (46%)	
L4: upper GI disease[†]		177 (8%)	
P: perianal[†]		563 (27%)	
E1: proctitis[b]			82 (8%)
E2: left-sided colitis[b]			357 (36%)
E3: extensive colitis[b]			558 (56%)
Pouch[†]	155 (5%)	38 (2%)	112 (9%)
Disease Activity at inclusion			
mHBI score[c]		1828 (100%)	
Remission 0-4		1218 (67%)	
Mild disease 5-7		314 (17%)	
Moderate disease 8-16		274 (15%)	
Severe disease > 16		22 (1.2%)	
mSCCAI score[d]			1016 (100%)
Remission < 2.5			752 (74%)
Active disease ≥ 2.5			264 (26%)
Liver disease due to IBD	3388 (100%)	2118 (100%)	1190 (100%)
Primary sclerosing cholangitis (PSC)[†]	71 (2%)	25 (1.2%)**	43 (4%)**
Liver disease other than PSC[†]	65 (1.9%)	42 (2.0%)	22 (1.8%)
Extraintestinal manifestations	3388 (100%)	2118 (100%)	1190 (100%)
Skin manifestations[†e]	336 (10%)	250 (12%)**	80 (7%)**
Musculoskeletal manifestations[†f]	731 (22%)	513 (24%)**	204 (17%)**
Ocular manifestations[†g]	147 (4%)	104 (5%)*	38 (3%)*
Complications	3388 (100%)	2118 (100%)	1190 (100%)
Osteopenia (T-score < -1)[†]	676 (20%)	496 (23%)**	169 (14%)**
Thromboembolic events[†]	119 (4%)	76 (4%)	42 (4%)

[a]Percentages calculated for 1677 patients with CD;
[b]Percentages calculated for 997 patients with UC;
[c]mHBI: modified Harvey-Bradshaw Index score; patients with CD were asked to rate their well-being on a scale from 1 to 10 (1: feeling terrible to 10: feeling very good) and to rate abdominal pain on a scale from 0 to 10 (0: no abdominal pain to 10: worst pain imaginable). Patients were also asked to provide data on diarrhoea frequency. In addition, patients were asked about the presence of oral aphthous lesions, active abscesses and fistulae as well as extraintestinal manifestations (arthralgia, uveitis, erythema nodosum, pyoderma gangrenosum). The physician assessed the presence of anal fissures and evaluated possible abdominal resistance through physical examination. mHBI data were available on 1828 patients (100%);
[d]mSCCAI: modified Simple Clinical Colitis Activity Index score; patients with UC were asked to rate their well-

being on a scale from 1 to 10 (1: feeling terrible to 10: feeling very good). In addition, patients were asked to describe the defecation frequency during the day and during the night, the defecation urgency (yes or no), the presence of blood in their stool (yes or no) and extracolonic manifestations (arthritis, uveitis, erythema nodosum, pyoderma gangrenosum);
[e]The following skin manifestations associated with IBD were scored: pyoderma gangrenosum, erythema nodosum, hidradenitis suppurativa, psoriasis or palmoplantar psoriasiform pustulosis and metastatic CD. Which type was not specified, only the presence of a skin manifestation;
[f]Musculoskeletal manifestations were divided into two groups: (1) arthritis (red and swollen joints), for example, dactylitis, reactive arthritis, gout; (2) arthropathy (not red or swollen joints, but symptoms with an inflammatory pattern; pain at night or at rest), for example, sacroiliitis, ankylosing spondylitis, enthesitis and inflammatory back pain;
[g]Ocular manifestations comprised uveitis and episcleritis diagnosed by a doctor. Which eye condition was not specified, only the presence of an ocular manifestation;
[†]Missing values were scored as non-present;
*P < 0.05; **P < 0.001.
CD: Crohn's disease; GI: gastrointestinal; IBD: inflammatory bowel disease (CD+UC+IBD-I (indeterminate)+IBD-U (unclassified)); UC: ulcerative colitis.

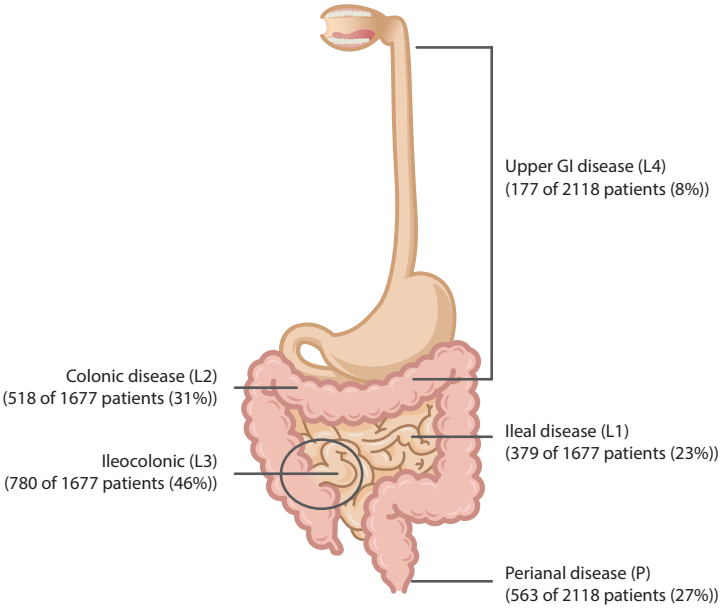


Figure 2 Disease localisation in patients with Crohn's disease in the Dutch IBD Biobank according to the Montreal classification.
GI: gastrointestinal.

Table 6 Malignancies, surgery and medication use of patients with inflammatory bowel disease in the Parelsnoer Institute cohort

	IBD	CD	UC
n (%)	3388 (100%)	2118 (62%)	1190 (35%)
Malignancy	3388	2118	1190
Dysplasia n[a]	131	62	63
Bowel cancer n[b]	15	9	5
Surgery	3388 (100%)	2118 (100%)	1190 (100%)
(Segmental) small bowel resection[†]	252 (7%)	242 (11%)	10 (0.8%)
Ileocaecal resection[†]	759 (22%)	758 (36%)	-
(Segmental) colon resection[†]	591 (17%)	368 (17%)	212 (18%)
Resection other[†]	168 (5%)	139 (7%)	28 (2%)
Strictureplasty[†]	99 (3%)	89 (4%)	-
Ileostomy/colostomy[†]	414 (12%)	283 (13%)	123 (10%)
Surgery for abscesses or fistulas[†]	494 (15%)	467 (22%)	27 (2%)
Outcome post-surgery	3388 (100%)	2118 (100%)	1190 (100%)
Stoma[†]	402 (12%)	270 (13%)	121 (10%)
Disease recurrence after IBD surgery			
Neoterminal ileum[c]	393 (52%%)	393 (52%)	-
Ileocolonic anastomosis[d]	56 (7%)	56 (7%)	-
Pouchitis[e]	93 (60%)	22 (58%)	67 (60%)
Surgical complication	1187 (100%)	959 (100%)	216 (100%)
Stricture anastomosis[f]	122 (10%)	107 (11%)	15 (7%)
Medication use during disease course	3306 (100%)	2068 (100%)	1158 (100%)
Immunomodulators[g]	2216 (67%)	1513 (73%)**	664 (57%)**
Biologicals[h]	1274 (39%)	1027 (50%)**	231 (20%)**
Azathioprine[i]	1374 (42%)	951 (46%)**	398 (34%)**
Mercaptopurine[j]	276 (8%)	199 (10%)**	73 (6%)**
Both azathioprine and mercaptopurine[k]	270 (8%)	172 (8%)	90 (8%)
Thioguanine[l]	114 (3%)	62 (3%)	50 (4%)
Methotrexate[m]	423 (13%)	363 (18%)**	52 (4%)**

[a]Dysplasia had to be confirmed in an intestinal biopsy by a pathologist. All intestinal biopsies were included including those from polyps;
[b]Bowel cancer included colorectal cancer, small bowel cancer and anal cancer;
[c]Percentage of disease recurrence in neoterminal ileum calculated from total patients with an ileocaecal resection (n = 759 IBD, n = 758 CD);
[d]Percentage of disease recurrence in ileocolonic anastomosis (no disease recurrence in neoterminal ileum) calculated from total patients with an ileocaecal resection (n = 759 IBD, n = 758 CD);
[e]Percentage of pouchitis calculated from total pouches (n = 155 IBD, n = 38 CD, n = 112 UC).
[f]Total patients who underwent surgery (small bowel resection, Ileocaecal resection, colon resection or resection other) (n = 1187 IBD, n = 959 CD, n = 216 UC);
[g]Immunomodulators: patients used one of the following immunosuppressives: azathioprine, Imuran, mercaptopurine, Purinethol, methotrexate, Metoject, thioguanine, Lanvis;
[h]Biologicals: patients used one of the following anti-tumour necrosis factors: infliximab, adalimumab or certolizumab;
[i]Azathioprine: patients used azathioprine or Imuran;
[j]Mercaptopurine: patients used mercaptopurine or Purinethol;
[k]Both azathioprine and mercaptopurine: patients used azathioprine and/or Imuran and mercaptopurine and/

or Purinethol. It was unclear which one of the drugs was used first;
[l]Thioguanine: patients used thioguanine or Lanvis;
[m]Methotrexate: patients used methotrexate or Metoject;
[†]Missing values were scored as non-present;
**P < 0.001.
CD: Crohn's disease; IBD: inflammatory bowel disease (CD+UC+IBD-I (indeterminate)+IBD-U (unclassified)); UC: ulcerative colitis.

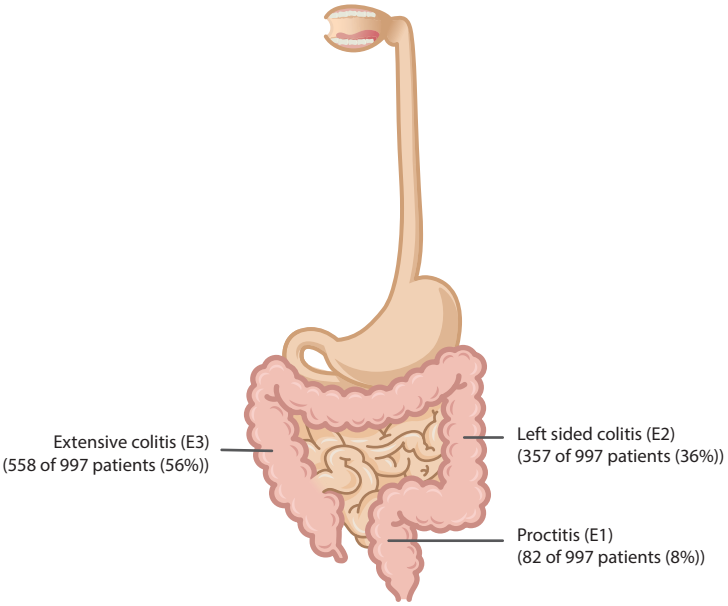


Figure 3 Disease localisation in patients with ulcerative colitis in the Dutch IBD Biobank according to the Montreal classification.

Genetic predictor of a fibrostenotic or inflammatory disease course in CD

The availability of genomic data and detailed clinical data in the Dutch IBD Biobank enabled a GWAS that aimed to find genetic predictors for recurrent fibrostenotic disease in patients with CD, by comparing the extremes of the clinical spectrum: (1) patients with CD with a mild disease course defined by inflammation without any signs of stricturing or penetrating disease during the last 5 years, versus (2) patients with CD who underwent ileocaecal resection due to confirmed intestinal strictures at least twice. We identified a genetic variant in the *WWOX* gene that regulates fibrosis through the SMAD pathway. The *WWOX* gene could therefore be an important signalling modulator involved in fibrostenotic CD (*Resubmitted to the Journal of Crohn's and Colitis*).

Previously published finding: Rare variants in MUC2 are associated with UC in the Dutch population. A subsequent study aimed to identify rare genetic variants with a large effect on UC susceptibility. Pooled resequencing of 122 genes in UC susceptibility loci in 1021 Dutch UC cases and 1166 Dutch controls revealed that rare variants in the *MUC2* gene were associated with increased

UC susceptibility (gene-based analysis with SKAT-O, nine variants in the *MUC2* gene: P value of 9.2×10^{-5}; threshold P = 0.0011 after Bonferroni correction). Interestingly, this association appeared to be population specific for the Netherlands.[25] Using the same approach and samples, a protein truncating variant in *RNF186* that protects against UC was also identified.[26]

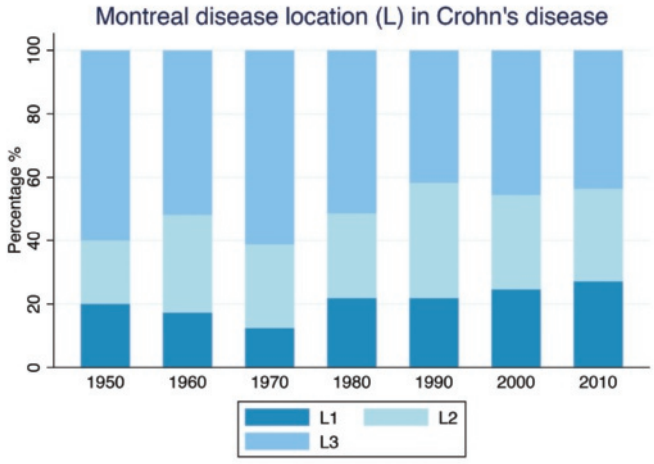


Figure 4 Date of Crohn's disease diagnosis and of disease location (L) according to the Montreal classification. L1: ileal; L2: colonic; L3: ileocolonic.

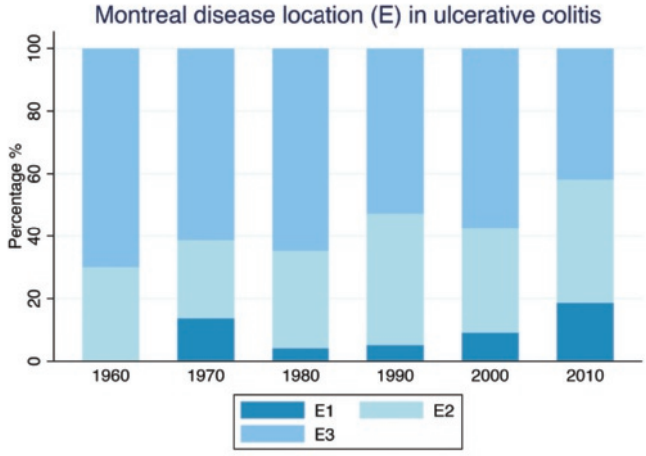


Figure 5 Date of ulcerative colitis diagnosis and of disease extent (E) according to the Montreal classification. E1: proctitis; E2: left-sided colitis; E3: pancolitis.

Associations between genetic variants and subphenotypes of IBD

The Dutch IBD Biobank participated in a large study where the clinical characteristics of patients with IBD were associated to genetic variants. The discovery of genetic variants associated with specific disease location and disease behaviour was published in the *Lancet*.[27]

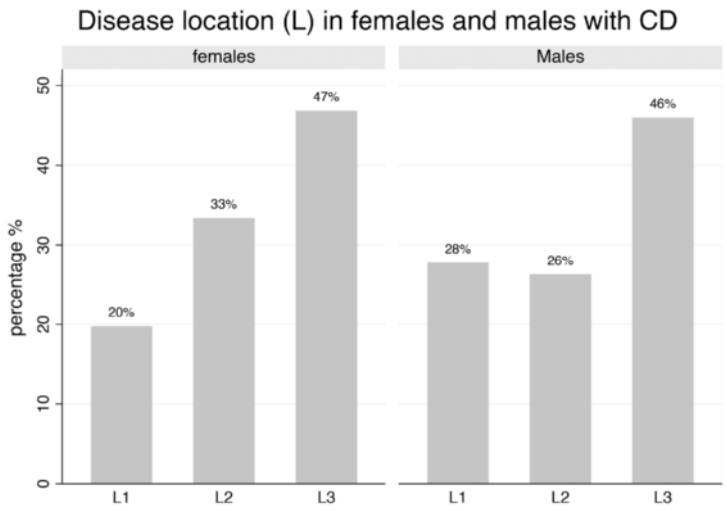


Figure 6 Disease location (L) according to the Montreal classification stratified by sex in patients with Crohn's disease (CD).
L1: ileal; L2: colonic; L3: ileocolonic.

GWAS and sequencing studies investigating the IBD diagnosis using DNA collections that were integrated in the Dutch IBD Biobank

For 1904 participants of the Dutch IBD Biobank genotype data are available consisting of ~200,000 single nucleotide polymorphisms (SNP) obtained using the Immunochip, an Illumina genotyping array focused on immune-mediated diseases. These genotype data were used in landmark genetic studies published in *Nature* and *Nature Genetics* investigating IBD pathogenesis.[3,28-30] These studies led to the discovery of 200 genetic loci associated with IBD, explaining 21.3% of the onset of IBD.

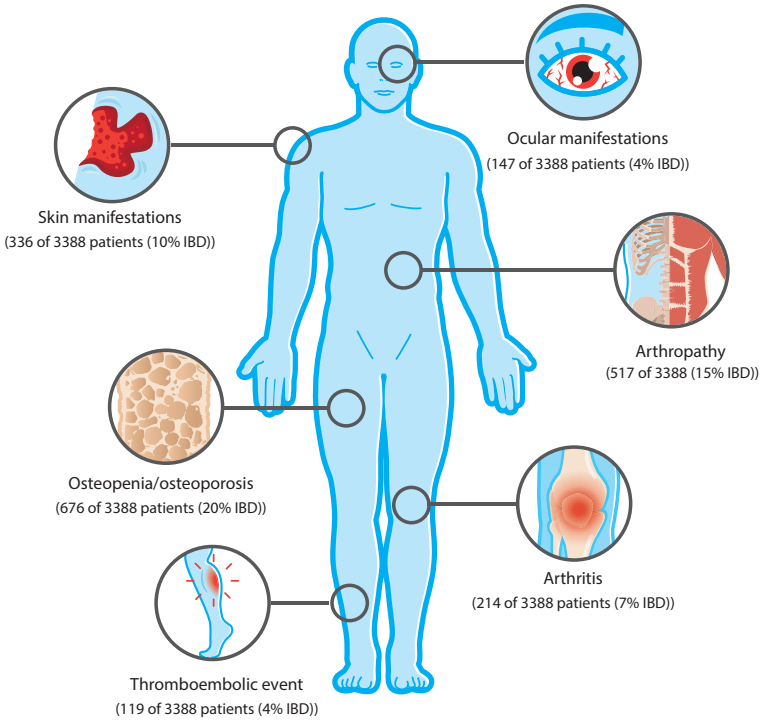


Figure 7 Extraintestinal manifestations and complications of patients with inflammatory bowel disease (IBD) in the Dutch IBD Biobank.

Discussion: strengths and weaknesses of the Dutch IBD Biobank

Strengths

A major strength of the Dutch IBD Biobank is its prospective design and extensive uniform information model comprising 225 data items, and the participation of all eight UMCs in the Netherlands. In addition, the biomaterials such as serum, DNA and a stool sample are collected at baseline, and, if available, biopsies from endoscopy and resection tissue are collected during follow-up, allowing the integration of subphenotypes enabling biomarker discovery research.

Since IBD is a chronic disease that requires lifelong treatment, patients treated in tertiary centres are rarely referred back to a general or local hospital and therefore loss to follow-up is uncommon.

Barriers to establishment and limitations

Setting up the Dutch IBD Biobank required a tremendous effort and there were many barriers to establishment. While some of these barriers were overcome, some limitations of the Dutch IBD Biobank remain. After a large initial grant provided by the Dutch government to the Netherlands Federation of University Medical Centres facilitating the establishment of the Dutch IBD Biobank and seven similar biobanks ended in 2011, the Dutch UMCs had to fund the continuation of the Dutch IBD Biobank themselves, meaning a reduction of staff that assisted in patient inclusion in some centres. As a consequence, the enrolment of patients has slowed down in these centres.

A major challenge was the establishment of the information technology (IT) infrastructure. In all UMCs, the local EHRs needed to be adapted so that the necessary information could be extracted. The gradual process of implementing data collection 'at the source' during the patient visit, and the renewal of EHRs in several hospitals means that adaptations to the local IT infrastructure continue to be necessary. Similar projects should be aware that the investments in the IT infrastructure will be ongoing after the establishment, and make sure they anticipate that continuous funding is required.

Data completeness, data similarity, data validation, quality control and feedback

A large majority of the data items were completely scored as can be seen in Tables 3-6. However, the different collection approaches by different UMCs sometimes lead to small differences in the clinical data, as some items were scored differently. Prior to completing this study, the authors reviewed all data and reported all inconsistencies to the national coordinators and to all UMCs. Several gastroenterologists, research nurses and IT departments improved the local data and a new upload to the Central Database was performed. Initially, very strict data validation steps were included in the Central Database software. However, these validation steps were too strict, and, because clinical patient records are often imperfect, very few patient records could be uploaded to the Central Database. After being aware of this problem, all data validation steps were removed from the Central Database software. Unfortunately, the lack of data validation steps led to errors in the data. Now, a small set of data validation protocols is in place. We recommend similar initiatives to start with simple data validation protocols and gradually expand these as the data quality and collection protocols improve.

Selection bias

Because all tertiary referral centres in the Netherlands participate in the Dutch IBD Biobank, the cohort will contain a large fraction of patients with IBD with a more severe disease course. This IBD cohort is not therefore suitable for studies that require a population-based cohort, for example, studies on the incidence and prevalence of IBD manifestations.

Collaboration

IBD researchers of the Dutch UMCs can access the Dutch IBD Biobank data and biomaterials after their research proposal has been approved by the Scientific Committee of the Dutch IBD Biobank. Other researchers can use the data and biomaterials of the Dutch IBD Biobank, but have to establish a cooperation with one or more Dutch UMCs.

Research proposal and application process

Research proposals can be submitted to the Scientific Committee and the Institutional Review Board. Proposals are judged against the following criteria:

1. It is reasonably plausible that the proposed research could lead to new insights.
2. The aims in the research proposal can be met using the proposed research methodology.
3. The proposed research is in concordance with the patient informed consent.
4. The proposed research will be conducted by people in institutes and facilities that are skilled and able to conduct the research.
5. The research proposal does not request more data and biomaterials than necessary.
6. The research proposal meets reasonable standards.
7. The proposed research does not unacceptably conflict or overlap with other research proposals.

After the Scientific Committee has approved a research proposal, the data manager will provide the pseudonymised research data in the web-based environment, and will facilitate the biomaterial delivery to the researcher. Applicants do not have to pay a fee.

The Dutch IBD Biobank can be contacted via email: IBDParel@umcg.nl. More information can also be found on the PSI website: www.parelsnoer.org. The Dutch IBD Biobank aims to cooperate with international IBD research groups. The information model and the list of biomaterials are publicly available and can be downloaded from the PSI website. The Dutch IBD Biobank encourages other biobanks to use the same information model and biomaterial collection standards to enable larger international studies on IBD and we encourage similar initiatives to contact us in an early stage.

Future developments

Genotyping the entire Dutch IBD Biobank

All DNA samples are in the process of being genotyped with a newly developed genome-wide genotyping array from Illumina, containing 750,000 SNPs. These data will be leveraged by

imputation against whole genome sequence data of 700 Dutch individuals studied in the Genome of the Netherlands project.[31] The availability of the genotype data will enable more genetic studies.

Web-based data access for researchers

The Dutch IBD Biobank is working on a multiomics data sharing portal called the *Molgenis Research IBD Portal*, based on Molgenis software.[32] This portal will make summary level statistics publicly available.

Mobile app for patients

The web-based follow-up of Patient-Reported Outcome Measurements including clinical disease activity scores is another project that the Dutch IBD Biobank is implementing. Patients will regularly fill in online questionnaires on disease activity, treatment response, quality of life and quality of care. Several UMCs are using the *app* My IBD Coach: http://www.sananet.nl/mijn-ibd-coach.html. The use of this *app* for IBD eHealth was extensively tested in a trial led by the MUMC, the Netherlands, where it was proven effective in reducing the number of hospital admissions.[33]

Conclusions

The Dutch UMCs have together created a biobank containing data and biomaterials of more than 3000 patients with IBD. The creation of the Dutch IBD Biobank took a very large multicentre multiyear effort, and new projects continue to improve the infrastructure and data collection. The main objective of the biobank is to facilitate the biomarker discovery. By now, studies using the Dutch IBD Biobank have led to the discovery a genetic predictor of a more severe disease course in patients with CD, showing that combining *-omics* data with prospectively collected clinical records can lead to useful results. Whether the standardising of patient data collection and during the patient visits and questionnaires online improves the clinical care of patients with IBD in the Netherlands is not yet known, but studies investigating the use of online disease activity scores and early detection of IBD exacerbations in the Netherlands are showing a reduction in hospitalisations.[33] We encourage researchers who want to establish similar biobanks to contact us, and to take our important recommendations, including the continuous IT funding, and the step-by-step implementation of data quality measures described in the discussion, into account.

Supplementary Data

Supplementary data are available at *BMJ Open* online.

References

1. Vektis. Insurance healthcare data in the Netherlands. https://www.vektis.nl
2. van den Heuvel TRA, Jeuring SFG, Zeegers MP, *et al*. Evolution of IBD incidence, disease phenotype, and mortality; 20 years of epidemiologic research in the Dutch population based IBDSL cohort. *Epidemiology* 2016.
3. Liu JZ, van Sommeren S, Huang H, *et al*. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet* 2015;47:979-86.
4. Ananthakrishnan AN. Epidemiology and risk factors for IBD. *Nat Rev Gastroenterol Hepatol* 2015;12:205-17.
5. Gevers D, Kugathasan S, Denson LA, *et al*. The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe* 2014;15:382-92.
6. Morgan XC, Tickle TL, Sokol H, *et al*. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol* 2012;13:R79.
7. Manniën J, Ledderhof T, Verspaget HW, *et al.* The Parelsnoer institute: a national network of standardized clinical biobanks in the Netherlands. *Open J Bioresour* 2017;25:1-8.
8. van Ommen GJ, Törnwall O, Bréchot C, *et al.* BBMRI-ERIC as a resource for pharmaceutical and life science industries: the development of biobank-based Expert Centres. *Eur J Hum Genet* 2015;23:893-900.
9. Spekhorst LM, Visschedijk MC, Alberts R, *et al.* Performance of the Montreal classification for inflammatory bowel diseases. *World J Gastroenterol* 2014;20:15374-81.
10. Centres NF of UM. Collecting data at the source' (Dutch title: Registratie aan de bron). Utrecht: NFU Nederlandse Federatie van Universitair Medische Centra.
11. NEN. Health Informatics - Information security management in healthcare. 7510:2011 NL. Vlinderweg, Delft, Netherlands: NEN.
12. International Organization for Standardization. Information technology - Security techniques - Information security management systems - Requirements. Geneva Switzerland: International Organization for Standardization; ISO/IEC 27001:2013.
13. Lennard-Jones JE. Classification of inflammatory bowel disease. *Scand J Gastroenterol* Suppl 1989;170:6-9.
14. Gasche C, Scholmerich J, Brynskov J, *et al.* A simple classification of Crohn's disease: report of the Working Party for the World Congresses of Gastroenterology, Vienna 1998. *Inflamm Bowel Dis* 2000;6:8-15.
15. Burisch J, Pedersen N, Čuković-Čavka S, et al. East-West gradient in the incidence of inflammatory bowel disease in Europe: the ECCO-EpiCom inception cohort. *Gut* 2014;63:588-97.
16. Sjöberg D, Holmström T, Larsson M, *et al.* Incidence and natural history of ulcerative colitis in the Uppsala Region of Sweden 2005-2009 - results from the IBD cohort of the Uppsala Region (ICURE). *J Crohns Colitis* 2013;7:e351-7.
17. Moum B, Vatn MH, Ekbom A, *et al*. Incidence of ulcerative colitis and indeterminate colitis in four counties of southeastern Norway, 1990-93. A prospective population-based study. The Inflammatory Bowel South-Eastern Norway (IBSEN) Study Group of Gastroenterologists. *Scand J Gastroenterol* 1996;31:362-6.
18. Lakatos L, Kiss LS, David G, *et al.* Incidence, disease phenotype at diagnosis, and early disease course in inflammatory bowel diseases in Western Hungary, 2002-2006. *Inflamm Bowel Dis* 2011;17:2558-65.
19. Isene R, Bernklev T, Høie O, *et al*. Extraintestinal manifestations in Crohn's disease and ulcerative colitis: results from a prospective, population-based European inception cohort. *Scand J Gastroenterol* 2015;50:300-5.
20. Veloso FT, Carvalho J, Magro F. Immune-related systemic manifestations of inflammatory bowel disease. A prospective study of 792 patients. *J Clin Gastroenterol* 1996;23:29-34.
21. Bernstein CN, Blanchard JF, Rawsthorne P, *et al.* The prevalence of extraintestinal diseases in inflammatory bowel disease: a population-based study. *Am J Gastroenterol* 2001;96:1116-22.
22. Manguso F, Sanges M, Staiano T, *et al*. Cigarette smoking and appendectomy are risk factors for extraintestinal manifestations in ulcerative colitis. *Am J Gastroenterol* 2004;99:327-34.

23. Roberts H, Rai SN, Pan J, *et al.* Extraintestinal manifestations of inflammatory bowel disease and the influence of smoking. *Digestion* 2014;90:122-9.

24. Ott C, Takses A, Obermeier F, *et al.* Smoking increases the risk of extraintestinal manifestations in Crohn's disease. *World J Gastroenterol* 2014;20:12269-76.

25. Visschedijk MC, Alberts R, Mucha S, *et al.* Pooled Resequencing of 122 Ulcerative Colitis Genes in a Large Dutch Cohort Suggests Population-Specific Associations of Rare Variants in MUC2. *PLoS One* 2016;11:e0159609.

26. Rivas MA, Graham D, Sulem P, *et al.* A protein-truncating R179X variant in RNF186 confers protection against ulcerative colitis. *Nat Commun* 2016;7:12342.

27. Cleynen I, Boucher G, Jostins L, *et al.* Inherited determinants of Crohn's disease and ulcerative colitis phenotypes: a genetic association study. *Lancet* 2016;387:156-67.

28. Franke A, McGovern DP, Barrett JC, *et al.* Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet* 2010;42:1118-25.

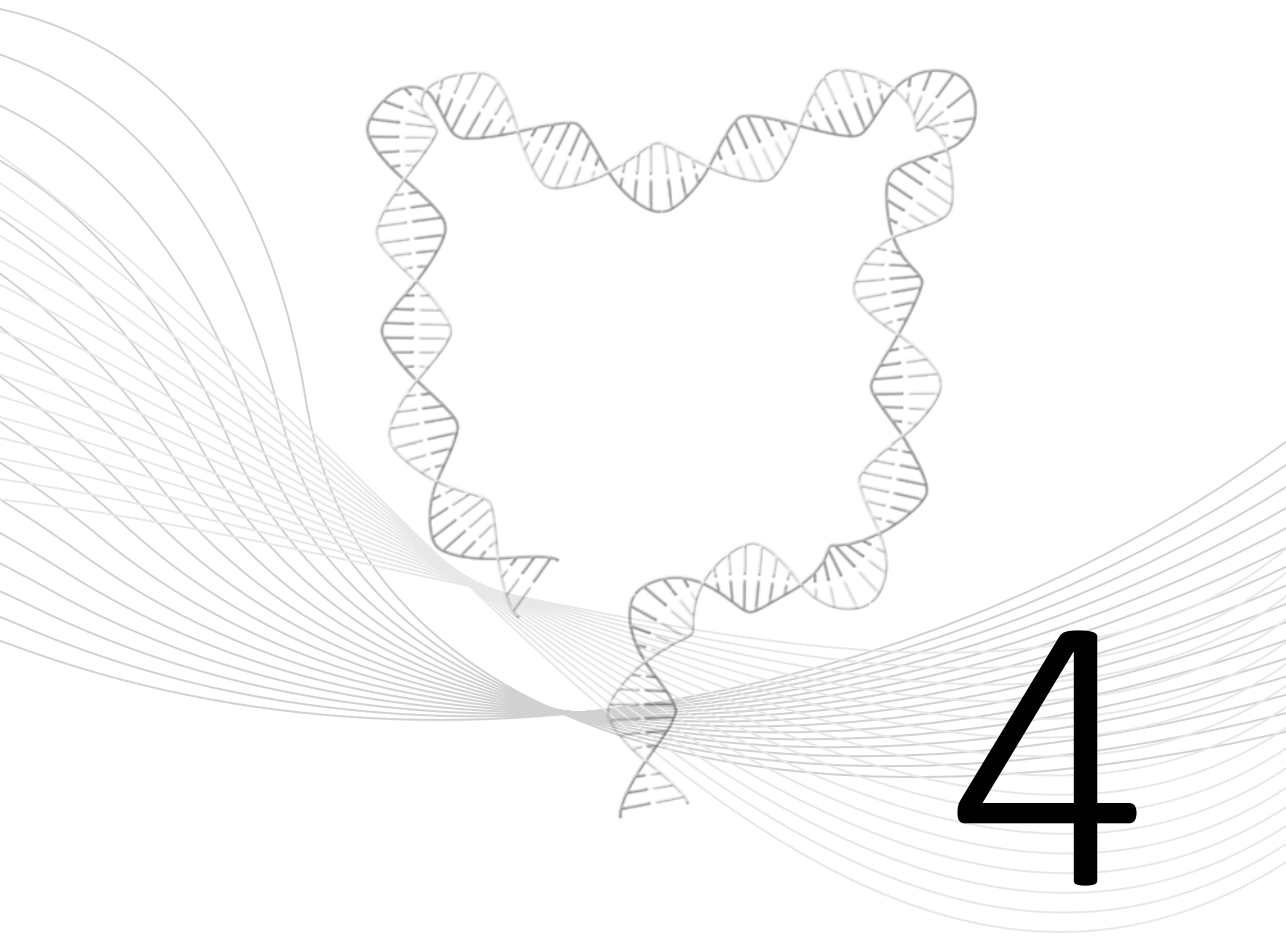
29. Jostins L, Ripke S, Weersma RK, *et al.* Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;491:119-24.

30. Anderson CA, Boucher G, Lees CW, *et al.* Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet* 2011;43:246-52.

31. Francioli LC, Menelaou A, Pulit SL, *et al.* Whole-genome sequence variation, population structure and demographic history of the Dutch population. *Nat Genet* 2014;46:818-25.

32. Swertz MA, Dijkstra M, Adamusiak T, *et al.* The MOLGENIS toolkit: rapid prototyping of biosoftware at the push of a button. *BMC Bioinformatics* 2010;11(Suppl 12):S12

33. de Jong MJ, van der Meulen-de Jong AE, Romberg-Camps MJ, *et al.* Telemedicine for management of inflammatory bowel disease (myIBDcoach): a pragmatic, multicentre, randomised controlled trial. *Lancet* 2017;390:959-68.



4

Prevalence of- and risk factors for work disability in Dutch patients with inflammatory bowel disease

LM Spekhorst, B Oldenburg, AA van Bodegraven, DJ de Jong, F Imhann,
AE van der Meulen-de Jong, MJ Pierik, JC van der Woude, G Dijkstra, G D'Haens,
M Löwenberg, RK Weersma, EAM Festen;
on behalf of the Parelsnoer Institute (PSI) and the Dutch Initiative on Crohn and Colitis (ICC)

World J Gastroenterol 2017 Dec 14;23(46):8182-8192

Abstract

Aim: To determine the prevalence of work disability in inflammatory bowel disease (IBD), and to assess risk factors associated with work disability.

Methods: For this retrospective cohort study, we retrieved clinical data from the Dutch IBD Biobank on July 2014, containing electronic patient records of 3388 IBD patients treated in the eight University Medical Centers in the Netherlands. Prevalence of work disability was assessed in 2794 IBD patients and compared with the general Dutch population. Multivariate analyses were performed for work disability (sick leave, partial and full disability) and long-term full work disability (> 80% work disability for > 2 years).

Results: Prevalence of work disability was higher in Crohn's disease (CD) (29%) and ulcerative colitis (UC) (19%) patients compared to the general Dutch population (7%). In all IBD patients, female sex, a lower education level, and extraintestinal manifestations, were associated with work disability. In CD patients, an age > 40 years at diagnosis, disease duration > 15 years, smoking, surgical interventions, and anti-TNFα use were associated with work disability. In UC patients, an age > 55 years, and immunomodulator use were associated with work disability. In CD patients, a lower education level (OR = 1.62, 95% CI: 1.02-2.58), and in UC patients, disease complications (OR = 3.39, 95% CI: 1.09-10.58) were associated with long-term full work disability.

Conclusion: The prevalence of work disability in IBD patients is higher than in the general Dutch population. Early assessment of risk factors for work disability is necessary, as work disability is substantial among IBD patients.

Introduction

Inflammatory bowel disease (IBD), consisting of Crohn's disease (CD) and ulcerative colitis (UC) is a heterogeneous/multifaceted disease. Patients suffer from symptoms like diarrhoea, abdominal pain, fatigue, and weight loss. The disease course is unpredictable, complicated by flares, need for chronic medication use and need for surgery. This leads to increasing intestinal damage with a high disease burden in many patients with IBD.[1] It has been established that, depending on disease activity, patients with IBD experience a lower quality of life.[2,3] IBD generally makes its debut during the second or third decade of life. Therefore, patients with IBD can encounter major problems during their economically productive life, which can lead to work disability, sometimes at young age. Work disability in patients with IBD is associated with a further decrease in quality of life, and high societal costs, especially if work disability is long-term and arises at a young age.[4,5] Preventing work disability is therefore an important goal in IBD management.

Work disability rates have been reported before, but reported rates are heterogeneous, ranging between 5% and 33% depending on sample size, methodology, and study population.[6-12] Several clinical factors have been suggested to play a role in work disability, such as female sex, age, disease duration, disease severity, and surgical interventions.[10-16] However, in most study designs, risk factors for any form of work disability are the focus of interest, neglecting risk factors associated specifically with long-term full work disability (> 80% work disability for > 2 years).

Therefore, our aims were: (1) to assess the prevalence of work disability in patients with IBD and to compare this with the general Dutch population; and (2) to identify risk factors for work disability, especially risk factors for long-term full work disability.

Materials and methods

Study design and study population

For this study, we used data from the Dutch IBD Biobank, which is part of the Parelsnoer Institute (www.parelsnoer.org). Since 2007, every patient with IBD that is treated in one of the eight Dutch University Medical Centers (UMCs) is invited to participate. The data is collected prospectively and patients with IBD are being enrolled continuously. Clinical data is retrieved from medical records by using a standardized information model containing 225 IBD related items. Items used for this study are; demographic items, diagnosis, smoking status, employment status, disease location, disease behaviour, education level, surgery-related items, medication use, extraintestinal manifestations and complications. Definitions of these items can be found in Table 1. At the moment of data freeze (the 17th of July 2014), the Dutch IBD Biobank contained 3388 patients with IBD. For this study,

patients for whom employment status was missing were excluded. We only included adult patients with IBD within working age, and thus excluded patients over the age of 65 years.

Table 1 Definition of used Parelsnoer Information Model items

Diagnosis of IBD	
Crohn's disease Ulcerative colitis IBD-unclassified IBD-Indeterminate	Diagnosis of IBD was defined by clinical symptoms of the disease and confirmed by endoscopy, radiology or histology. If differentiation between CD and UC was not possible, patient were classified as Inflammatory Bowel Disease Unclassified (IBD-U). In the case that a pathologist was not able to differentiate between CD or UC following colectomy, the patient was classified as Inflammatory Bowel Disease Indeterminate (IBD-I).
Family history of IBD	Family history of IBD was registered up to 3rd degree relatives
Smoking status	Smoking status at diagnosis was documented. Patients were classified as never smokers, current smokers or former smokers
Education level	Patients were classified in two groups; low education (Lower general education; Lower vocational education; General secondary education; Vocational secondary education; Did not finish primary school) and high education (Pre-university secondary education; Vocational post-secondary education; University)
Disease behaviour	Disease behaviour is reported if it occurred at any point in the disease course up to baseline.
Fistulising disease	Fistulising disease is confirmed by physical examination, radiological or endoscopy assessment.
Stricturing disease	A stricture had to be symptomatic. An anal stenosis had to be symptomatic and confirmed by physical examination.
Penetrating disease	Penetrating disease is defined as an intestinal abscess or intestinal perforation due to disease activity.
Prescribed medication	
Anti-tumour necrosis factor alpha	Prescribed medication included anti-tumour necrosis factor alpha (TNF-alpha) agents (infliximab, adalimumab or certolizumab). Data on medication use was available for the entire disease course; medication use was therefore defined as medication 'ever used'
Immunomodulators	immunomodulators (mercaptopurine, azathioprine, thioguanine or methotrexate). Data on medication use was available for the entire disease course; medication use was therefore defined as medication 'ever used'
IBD-related surgical interventions	IBD-related surgical interventions included small bowel resection, ileocaecal resection, colon resection, resection other, strictureplasty, ileostomy/ colostomy, surgery for abscesses or fistula, stoma and pouch
Extraintestinal manifestations	
Skin manifestations	Skin manifestations unrelated to the use of IBD medication such as pyoderma gangrenosum, erythema nodosum, hidradenitis suppurativa, psoriasis or palmoplantar psoriasiform pustolosis, and metastatic CD.

Table 1 *continued*

Extraintestinal manifestations	
Musculoskeletal manifestations	Arthritis was defined as red and swollen joints, dactylitis, reactive arthritis and gout. Arthropathy was defined as painful joints, without any swelling or redness, with an inflammatory pattern: pain at night or at rest (e.g. sacroiliitis, ankylosing spondylitis, enthesitis, and inflammatory back pain).
Ocular manifestations	Ocular manifestations included uveitis and episcleritis.
Complications	
Osteopenia	Osteopenia was defined as a bone mineral density T-score lower than -1.
Thromboembolic event	A thromboembolic event was confirmed by additional tests (radiology, endoscopy or histology).

IBD: Inflammatory bowel disease; CD: Crohn's disease; UC: ulcerative colitis.

Definition of work disability

Patients were asked about their employment status at inclusion, which was defined using the following categories; (1) working full time (80%-100%, or working 30 h or more a week); (2) working part time (< 80%, or working less than 30 h a week); (3) partial work disability (35%-80% work disability for > 2 years); (4) full work disability (> 80% work disability for > 2 years); (5) sick leave (work disability for < 2 years); (6) retired; and (7) "Other".

A Dutch law ("Wet werk en inkomen naar arbeidsvermogen" or WIA) prescribes that from the moment Dutch citizens are disabled to work they are entitled to receive a maximum of 170% of their wages during a period of 104 mo. (generally 100% for 52 wk., followed by 70% for 52 wk.). If after 104 mo. the person is still on sick leave the same law prescribes whether patients receive a disability pension. There are two types of disability pension; partial disability (35%-80% work disability) and full disability (more than 80% work disability). A specialized physician determines the percentage of work disability based on questionnaires and physical examination.

Due to the nature of our data we could not distinguish whether work disability was solely attributable to IBD or whether there was an additional cause.

To compare the currently reported work disability rates with the general Dutch population we retrieved information from Statistics Netherlands (CBS), a national institute gathering statistical information about the Dutch population, including employment rates.[17,18] Employment rates were collected for 2014 and matched for sex and age.

Statistical analysis

In all analyses, patients with UC were grouped with patients with IBD-Unclassified (IBD-U) and IBD-Indeterminate (IBD-I). Dichotomous variables were compared using the Pearson $\chi 2$ test. Continuous variables were compared using the Mann-Whitney-*U* test. The Student *t*-test was used

to compare overall work disability in our cohort to the general Dutch population. The variable "age" represented the age at inclusion at which employment status was assessed. To assess clinical predictors of work disability in patients with IBD, we categorized the variable employment status into two groups; employed (both full time and part time) and disabled for work (including partial disability, full disability and sick leave). Retirement and "other" were left out of the final analyses after we determined that "other" was unlikely to contain undeclared differential work disability, such as pregnancy leave. We performed a multivariate analysis with employment status as outcome and included demographic and clinical items that had a $P < 0.10$ in the univariate analysis for employment status. Analyses were repeated in patients with CD and patients with UC, separately. To assess risk factors for long-term full work disability, patients with full disability were compared to patients with partial disability. Statistical analyses were performed with Stata Software V.13.1.[19]

Results

Patient population and demographic characteristics

A total of 2794 patients with IBD were within working age and were included in the analysis; 1740 patients with CD and 1054 patients with UC. Demographic and clinical characteristics are depicted in Table 2. In our cohort, more patients with CD were female compared to patients with UC (63% vs 53%, $P < 0.01$). Age at diagnosis was lower in patients with CD compared to patients with UC (24 years old vs 29 years old, $P < 0.01$). The prevalence of a family history positive for IBD (30% vs 25%, $P = 0.01$) as well as the appendectomy rate (15% vs 7%, $P < 0.01$) were higher in patients with CD. Patients with CD were more often smokers than patients with UC (27% vs 12%, $P < 0.01$). According to the Montreal classification, most patients with CD had ileocolonic disease (47% (L3)), whereas most patients with UC had an extensive colitis (58% (E3)) (Table 2).

Employment status in patients with IBD compared to the general Dutch population

Figure 1 shows the percentages of work disability in patients with CD, patients with UC and the general Dutch population per age category in females (Figure 1A) and males (Figure 1B). Overall, work disability was significantly higher in IBD patients compared to the general Dutch population, independent of sex ($P < 0.05$).

Table 2 Demographic and clinical characteristics in patients with Crohn's disease and ulcerative colitis *n* (%)

	CD	UC	P value
n	1740	1054	
Sex (female)	1103 (63%)	561 (53%)	P < 0.01
Median Age years (IQR 25-75)	39.9 (30-51)	43.4 (33-53)	P < 0.01
Median Age Diagnosis years (IQR 25-75)	24.1 (19-32)	29.0 (22-39)	P < 0.01
Median Age Disease duration years (IQR 25-75)	11.7 (6-21)	10.3 (5-18)	P < 0.01
Family history of IBD[†]	518 (30%)	266 (25%)	P = 0.01
Appendectomy[†]	266 (15%)	70 (7%)	P < 0.01
Smoking	1620 (100%)	984 (100%)	
Yes	439 (27%)	117 (12%)	P < 0.01
Disease location			
L1: ileal disease[a]	324 (22%)		
L2: colon disease[a]	450 (31%)		
L3: ileocolon disease[a]	696 (47%)		
L4: upper GI disease[†]	156 (9%)		
P: peri-anal disease[†]	485 (28%)		
E1: proctitis[b]		77 (8%)	
E2: left-sided colitis[b]		315 (34%)	
E3: extensive colitis[b]		541 (58%)	
Disease behaviour CD			
Fistulising disease[†]	289 (17%)		
Penetrating disease[†]	232 (13%)		
Stricturing disease[†]	445 (26%)		
Education	1708 (100%)	1036 (100%)	
Low	932 (55%)	517 (50%)	P = 0.02
Employment status			
Full-time	607 (35%)	440 (42%)	P < 0.01
Part-time	359 (21%)	224 (21%)	
Partially disabled to work	129 (7%)	54 (5%)	
Fully disabled to work	342 (20%)	124 (12%)	
Sick leave	77 (4%)	36 (3%)	
Retired	22 (1.3%)	21 (2%)	
Other	204 (12%)	155 (15%)	

[a]These percentages are calculated for 1470 patients with CD;
[b]These percentages are calculated for 993 patients with UC;
[†]Missing values were scored as non-present.
CD: Crohn's disease; UC: ulcerative colitis.

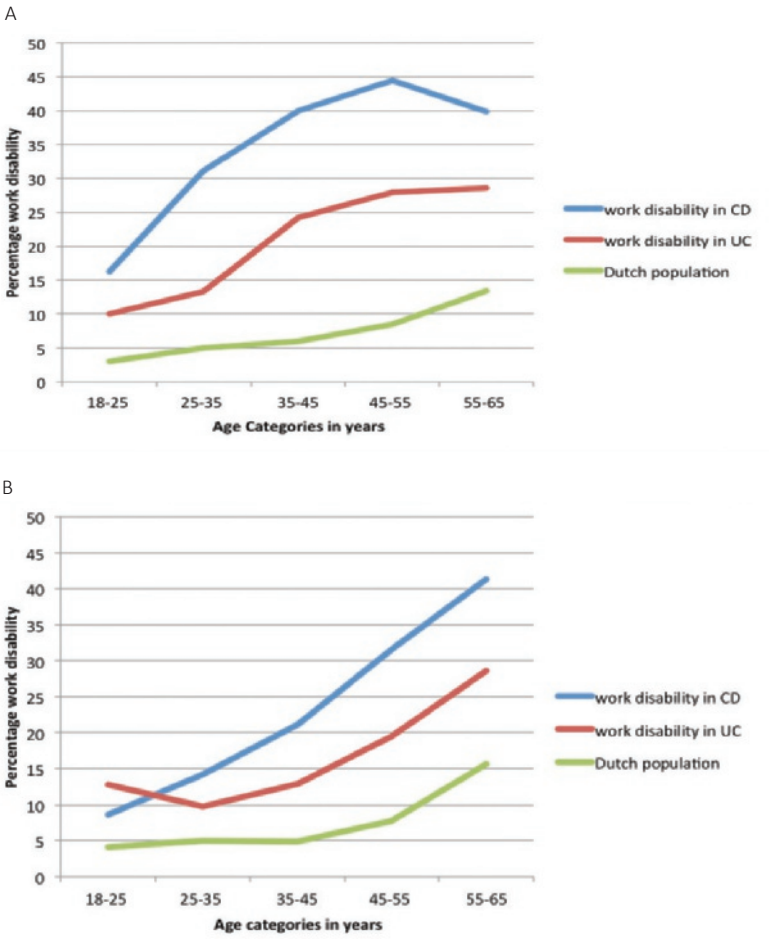


Figure 1 Percentage of work disability per age category in females and males. Comparing work disability rates in patients with CD and patients with UC with the general Dutch population. A: Work disability in females; B: Work disability in males.
CD: Crohn's disease; UC: ulcerative colitis.

Clinical risk factors of (long-term full) work disability in patients with CD

Table 3 depicts the (clinical) risk factors that were significantly associated with work disability in CD (548 patients with CD were fully disabled and 966 patients with CD were employed). In the multivariate model, female sex (OR = 2.03, 95% CI: 1.53-2.69), an age > 40 years at diagnosis (OR = 3.69, 95% CI: 1.83-7.42), a disease duration > 15 years (OR = 1.67, 95% CI: 1.15-2.43), a lower education level (OR = 2.02, 95% CI: 1.55-2.64), smoking (OR = 1.45, 95% CI: 1.09-1.92), surgical interventions (OR = 1.48, 95% CI: 1.10-2.00), anti-tumour necrosis factor alpha (TNFα) use (OR =

1.86, 95% CI: 1.43-2.43), and extraintestinal manifestations (OR = 1.36, 95% CI: 1.05-1.77) were all associated with an increased odds ratio for work disability in CD.

In the second analysis, partial disability was compared to long-term full work disability for CD. In the multivariate model, an age > 55 years (OR = 3.06, 95% CI: 1.54-6.07), and a lower education level (OR = 1.62, 95% CI: 1.02-2.58) were associated with long-term full work disability in CD (Table 4).

Clinical risk factors of (long-term full) work disability in patients with UC

In Table 5, (clinical) risk factors that were significantly associated with work disability in UC (214 patients with UC were fully disabled and 664 patients with UC were employed) have been shown. In the multivariate model, female sex (OR = 1.76, 95% CI: 1.23-2.53), an age > 55 years (OR = 2.93, 95% CI: 1.68-5.14), a lower education level (OR = 2.59, 95% CI: 1.81-3.70), immunomodulator use (OR = 1.58, 95% CI: 1.09-2.28), and extraintestinal manifestations (OR = 2.13, 95% CI: 1.47-3.09) were all associated with an increased odds ratio for work disability in UC.

In the second analysis, partial disability was compared to long-term full work disability for UC. In the multivariate model, an age > 55 years (OR = 3.49, 95% CI: 1.23-9.92), and complications (OR = 3.39, 95% CI: 1.09-10.58) were associated with long-term full work disability in UC (Table 6).

Table 3 Univariate and multivariate regression analyses of work disability in patients with Crohn's disease *n* (%)

	Work disability	Employed	Unadj. OR (95% CI)	Adj. OR (95% CI)
n (%)	548 (100%)	966 (100%)		
Sex (female)	397 (72%)	556 (58%)	1.94 (1.55-2.43)	2.03 (1.53-2.69)
Age				
< 40 years	215 (39%)	537 (55%)	1.00	1.00
40-55 years	220 (40%)	335 (35%)	1.64 (1.30-2.07)	0.94 (0.65-1.35)
> 55 years	113 (21%)	94 (10%)	3.00 (2.19-4.12)	1.63 (0.95-2.77)
Age at diagnosis				
A1: diagnosis ≤ 16 years	47 (9%)	156 (16%)	1.00	1.00
A2: diagnosis 17-40 years	413 (75%)	719 (75%)	1.91 (1.35-2.70)	2.02 (1.30-3.13)
A3: diagnosis > 40 years	88 (16%)	91 (9%)	3.21 (2.07-4.98)	3.69 (1.83-7.42)
Disease duration				
≤ 15 years	285 (52%)	618 (64%)	1.00	1.00
> 15 years	263 (48%)	348 (36%)	1.64 (1.32-2.03)	1.67 (1.15-2.43)
Education	537 (100%)	950 (100%)		
Low	366 (68%)	446 (47%)	2.42 (1.94-3.02)	2.02 (1.55-2.64)
Smoking	508 (100%)	904 (100%)		
Yes	180 (35%)	203 (22%)	1.90 (1.49-2.41)	1.45 (1.09-1.92)
Disease location				
L1: ileal disease[a]	88 (19%)	184 (23%)	1.00	1.00
L2: colon disease[a]	138 (30%)	263 (33%)	1.10 (0.79-1.52)	1.10 (0.75-1.61)
L3: ileocolon disease[a]	237 (51%)	359 (44%)	1.38 (1.02-1.87)	1.29 (0.91-1.83)
L4: upper GI disease[†]	52 (9%)	82 (8%)	1.13 (0.79-1.63)	-
P: peri-anal disease[†]	170 (31%)	256 (26%)	1.25 (0.99-1.57)	1.03 (0.77-1.37)
Disease behaviour CD				
Fistulising disease[†]	95 (17%)	156 (16%)	1.09 (0.82-1.44)	-
Stricturing disease[†]	145 (26%)	248 (26%)	1.04 (0.82-1.32)	-
Penetrating disease[†]	72 (13%)	128 (13%)	0.99 (0.73-1.35)	-
Pouch[†]	12 (2%)	17 (1.8%)	1.25 (0.59-2.64)	-
Stoma[†]	96 (18%)	108 (11%)	1.69 (1.25-2.27)	1.15 (0.78-1.69)
Surgery[†]	288 (53%)	406 (42%)	1.53 (1.24-1.89)	1.48 (1.10-2.00)
Medication	537 (100%)	949 (100%)		
Anti-TNFα	309 (58%)	435 (46%)	1.60 (1.29-1.98)	1.86 (1.43-2.43)
Immunomodulators	395 (74%)	698 (74%)	1.00 (0.79-1.27)	-
Extraintestinal manifestations[†]	229 (42%)	294 (30%)	1.64 (1.32-2.04)	1.36 (1.05-1.77)
PSC[†]	7 (1.3%)	14 (1.5%)	0.88 (0.35-2.19)	-
Complications[†]	155 (28%)	223 (23%)	1.31 (1.04-1.67)	1.18 (0.88-1.59)

[a]These percentages are calculated for 463 patients with CD that were disabled for work and 806 patients with CD that were employed;
[†]Missing values were scored as non-present.
CD: Crohn's disease; Unadj: Unadjusted; TNFα: Tumour necrosis factor alpha; PSC: Primary sclerosing cholangitis.

Table 4 Univariate and multivariate regression analyses of full and partial work disability in patients with Crohn's disease *n* (%)

	Full disability	Partial disability	Unadj. OR (95% CI)	Adj. OR (95% CI)
n (%)	342 (100%)	129 (100%)		
Sex (female)	251 (73%)	100 (78%)	0.80 (0.50-1.29)	-
Age				
< 40 years	110 (32%)	54 (42%)	1.00	1.00
40-55 years	137 (40%)	61 (47%)	1.10 (0.71-1.72)	1.05 (0.65-1.70)
> 55 years	95 (28%)	14 (11%)	3.33 (1.74-6.37)	3.06 (1.54-6.07)
Age at diagnosis				
A1: diagnosis ≤ 16 years	28 (8%)	13 (10%)	1.00	-
A2: diagnosis 17-40 years	256 (75%)	103 (80%)	1.15 (0.58-2.32)	
A3: diagnosis > 40 years	58 (17%)	13 (10%)	2.07 (0.85-5.05)	
Disease duration				
≤ 15 years	151 (44%)	65 (50%)	1.00	-
> 15 years	191 (56%)	64 (50%)	1.28 (0.86-1.93)	
Education	332 (100%)	128 (100%)		
Low	236 (71%)	73 (57%)	1.85 (1.21-2.83)	1.62 (1.02-2.58)
Smoking	323 (100%)	115 (100%)		
Yes	125 (39%)	33 (29%)	1.57 (0.99-2.49)	1.53 (0.93-2.50)
Disease location				
L1: ileal disease[a]	56 (19%)	21 (19%)	1.00	-
L2: colon disease[a]	80 (28%)	38 (35%)	0.79 (0.42-1.49)	
L3: ileocolon disease[a]	155 (53%)	50 (46%)	1.16 (0.64-2.11)	
L4: upper GI disease[†]	38 (11%)	8 (6%)	1.89 (0.86-4.17)	-
P: peri-anal disease[†]	109 (32%)	45 (35%)	0.87 (0.57-1.34)	-
Disease behaviour CD				
Fistulising disease[†]	64 (19%)	23 (18%)	1.06 (0.63-1.80)	-
Stricturing disease[†]	92 (27%)	33 (26%)	1.07 (0.67-1.70)	-
Penetrating disease[†]	47 (14%)	19 (15%)	0.92 (0.52-1.64)	-
Pouch[†]	6 (1.8%)	4 (3%)	0.56 (0.15-2.01)	-
Stoma[†]	69 (20%)	20 (16%)	1.38 (0.80-2.38)	-
Surgery[†]	192 (56%)	64 (50%)	1.30 (0.87-1.95)	-
Medication	335 (100%)	125 (100%)		
Anti-TNFα	193 (58%)	68 (54%)	1.14 (0.75-1.72)	-
Immunomodulators	249 (74%)	87 (70%)	1.26 (0.80-1.99)	-
Extraintestinal manifestations[†]	152 (44%)	50 (39%)	1.26 (0.84-1.91)	-
PSC[†]	2 (0.6%)	4 (3%)	0.18 (0.03-1.02)	0.19 (0.03-1.10)
Complications[†]	106 (31%)	34 (26%)	1.25 (0.80-1.98)	-

[a]These percentages are calculated for 291 patients with CD that were fully disabled and 109 patients with CD that were partially disabled;
[†]Missing values were scored as non-present.
CD: Crohn's disease; Unadj: Unadjusted; TNFα: Tumour necrosis factor alpha; PSC: Primary sclerosing cholangitis.

Table 5 Univariate and multivariate regression analyses of work disability in patients with ulcerative colitis *n* (%)

	Work disability	Employed	Unadj. OR (95% CI)	Adj. OR (95% CI)
n (%)	214 (100%)	664 (100%)		
Sex (female)	124 (58%)	333 (50%)	1.37 (1.00-1.87)	1.76 (1.23-2.53)
Age				
< 40 years	58 (27%)	287 (43%)	1.00	1.00
40-55 years	94 (44%)	277 (42%)	1.68 (1.16-2.42)	1.52 (0.97-2.40)
> 55 years	62 (29%)	100 (15%)	3.07 (2.01-4.69)	2.93 (1.68-5.14)
Age at diagnosis				
A1: diagnosis ≤ 16 years	18 (8%)	64 (10%)	1.00	1.00
A2: diagnosis 17-40 years	124 (58%)	474 (71%)	0.93 (0.53-1.63)	1.06 (0.55-2.04)
A3: diagnosis > 40 years	72 (34%)	126 (19%)	2.03 (1.12-3.69)	1.99 (0.92-4.28)
Disease duration				
≤ 15 years	129 (60%)	439 (66%)	1.00	-
> 15 years	85 (40%)	225 (34%)	1.29 (0.94-1.77)	
Education	208 (100%)	656 (100%)		
Low	144 (69%)	282 (43%)	2.98 (2.14-4.16)	2.59 (1.81-3.70)
Smoking	206 (100%)	621 (100%)		
Yes	30 (15%)	71 (11%)	1.32 (0.83-2.09)	-
Disease location				
E1: proctitis[a]	14 (7%)	54 (9%)	1.00	-
E2: left-sided colitis[a]	54 (27%)	215 (37%)	0.97 (0.50-1.87)	
E3: extensive colitis[a]	130 (66%)	317 (54%)	1.58 (0.85-2.95)	
Pouch[†]	37 (17%)	54 (8%)	2.36 (1.50-3.71)	1.72 (0.81-3.65)
Stoma[†]	35 (16%)	54 (8%)	2.21 (1.40-3.49)	1.13 (0.54-2.38)
Surgery[†]	60 (28%)	97 (15%)	2.28 (1.58-3.29)	1.68 (0.75-3.75)
Medication	211 (100%)	649 (100%)		
Anti-TNFα	53 (25%)	120 (18%)	1.48 (1.02-2.14)	1.33 (0.86-2.06)
Immunomodulators	139 (66%)	350 (54%)	1.65 (1.19-2.28)	1.58 (1.09-2.28)
Extraintestinal manifestations[†]	78 (36%)	130 (20%)	2.36 (1.68-3.30)	2.13 (1.47-3.09)
PSC[†]	8 (4%)	21 (3%)	1.19 (0.52-2.73)	-
Complications[†]	37 (17%)	89 (13%)	1.35 (0.89-2.05)	-

[a]These percentages are calculated for 198 patients with UC that were disabled for work and 586 patients with UC that were employed;
[†]Missing values were scored as non-present.
UC: ulcerative colitis; Unadj: Unadjusted; Adj: Adjusted; TNFα: Tumour necrosis factor alpha; PSC: Primary sclerosing cholangitis.

Table 6 Univariate and multivariate regression analyses of full and partial work disability in patients with ulcerative colitis *n* (%)

	Full disability	Partial disability	Unadj. OR (95% CI)	Adj. OR (95% CI)
n (%)	124 (100%)	54 (100%)		
Sex (female)	75 (60%)	26 (48%)	1.65 (0.87-3.14)	-
Age				
< 40 years	25 (20%)	15 (28%)	1.00	1.00
40-55 years	52 (42%)	32 (59%)	0.98 (0.45-2.12)	0.90 (0.40-2.03)
> 55 years	47 (38%)	7 (13%)	4.03 (1.45-11.17)	3.49 (1.23-9.92)
Age at diagnosis				
A1: diagnosis ≤ 16 years	7 (6%)	5 (9%)	1.00	-
A2: diagnosis 17-40 years	70 (56%)	34 (63%)	1.47 (0.43-4.97)	
A3: diagnosis > 40 years	47 (38%)	15 (28%)	2.24 (0.62-8.10)	
Disease duration				
≤ 15 years	65 (52%)	36 (67%)	1.00	1.00
> 15 years	59 (48%)	18 (33%)	1.82 (0.93-3.54)	1.78 (0.87-3.62)
Education	122 (100%)	50 (100%)		
Low	84 (69%)	37 (74%)	0.78 (0.37-1.63)	-
Smoking	121 (100%)	52 (100%)		
Yes	18 (15%)	7 (13%)	1.12 (0.44-2.88)	-
Disease location				
E1: proctitis[a]	9 (8%)	4 (8%)	1.00	-
E2: left-sided colitis[a]	39 (33%)	9 (18%)	1.93 (0.48-7.68)	-
E3: extensive colitis[a]	70 (59%)	38 (74%)	0.82 (0.24-2.84)	-
Pouch[†]	24 (19%)	9 (17%)	1.20 (0.52-2.79)	-
Stoma[†]	23 (19%)	8 (15%)	1.31 (0.54-3.15)	-
Surgery[†]	37 (30%)	14 (26%)	1.22 (0.59-2.50)	-
Medication	122 (100%)	53 (100%)		
Anti-TNFα	30 (25%)	10 (19%)	1.40 (0.63-3.13)	-
Immunomodulators	77 (63%)	37 (70%)	0.74 (0.37-1.48)	-
Extraintestinal manifestations[†]	51 (41%)	19 (35%)	1.29 (0.66-2.50)	-
PSC[†]	3 (2%)	4 (7%)	0.31 (0.07-1.44)	-
Complications[†]	28 (23%)	4 (7%)	3.65 (1.21-10.97)	3.39 (1.09-10.58)

[a]These percentages are calculated for 118 patients with UC that were fully disabled and 51 patients with UC that were partially disabled;
[†]Missing values were scored as non-present.
UC: ulcerative colitis; Unadj: Unadjusted; Adj: Adjusted; TNFα: Tumour necrosis factor alpha; PSC: Primary sclerosing cholangitis.

Discussion

In this nationwide clinical database and biobank study we assessed work disability rates in patients with IBD. We found higher rates of work disability in patients with IBD compared to the general Dutch population. Furthermore, in patients with IBD female sex, a lower education level, and extraintestinal manifestations were significantly associated with work disability. In patients with CD, an age > 40 years at diagnosis, disease duration > 15 years, smoking, surgical interventions, and anti-TNFα use were significantly associated with higher work disability. In patients with UC, an age > 55 years, and immunomodulator use were significantly associated with higher work disability. In patients with CD a lower education level was associated with long-term full work disability, whereas in patients with UC disease complications were associated with long-term full work disability.

High work disability rates in patients with IBD

In patients with IBD, the overall proportion of work disability was 29% in CD, and 19% in UC. These work disability rates were higher than those in the age-adjusted general Dutch population (7%). Suffering from (severe) IBD may in itself be explanatory for this higher percentage. Furthermore, all patients in the Dutch IBD Biobank were treated in tertiary referral centers with a severe disease course signature: 47% ileocolonic disease (L3 Montreal CD) and 58% extensive colitis (E3 Montreal UC). Indeed, disease in remission was associated with increased employment in CD[5] whereas disease severity was a predictor of work disability in IBD.[16]

Criteria for disability pension differ between countries due to political and socioeconomic factors. Comparison of work disability rates between different countries is therefore difficult. In addition, differences in selection criteria, study methodology, sample size, and definitions of work disability result in highly variable disability rates reported (from 5% to 33%).[6-14] In two comparable Dutch IBD-population studies it has been concluded by the authors that disability rates in CD and in UC were higher compared to the general Dutch population, which is in line with the current results. However, work disability rates of 33% in CD, 24% in UC, and 11% in the control population were reported in the first study that was conducted in 2002,[8] which is higher than the disability rates we found. This may be due to changes in the welfare system since 2002, when patients got a disability pension after one year of sick leave, instead of after two years as it is now. Disability rates in a second, more recent Dutch study, by van der Valk *et al*,[10] were more in line with our findings but lower (18% in CD, 10% in UC, and 7% in the control population). Although there seems to be a wide range in work disability rates in IBD, it may be concluded that the rate of work disability is substantial among patients with IBD.

Clinical risk factor for work disability

Female sex,[10,11,13,15,16] an age > 40 years at diagnosis,[11,13] older age at inclusion,[10,12,13,15,16] a lower education level,[10,13] extraintestinal manifestations,[10,12,16] a disease duration > 15 years,[12,15,16] surgical interventions,[11-13] and immunomodulator use[16] have been reported to be associated with work disability in patients with IBD.

In UC, an age > 55 years was associated with work disability, which might not be directly disease related. Indeed, it has been reported that disability pensions were not caused by UC but often by other age-related comorbidities.[11,12] On the other hand, in CD it has been found that patients with CD receive their disability pension because of IBD or other IBD related comorbidity,[11] rather than other age related comorbidities.

Surgical interventions were a risk factor for work disability in patients with CD, but not in UC. As surgery is an indicator of severe disease these data seemed to corroborate the hypothesis that severity of disease predicts work disability. In line with this, higher work disability rates have been reported after colectomy in patients with UC.[14] A recent review has shown that surgical rates have declined since the introduction of biologicals, indicating a beneficial effect of treatment with biologicals on disease outcome.[20] Our study is not well suited to study this effect, since our most reliable clinical information has been collected after the introduction of biologicals. Other studies report no differences in disability between patients receiving anti-TNFα treatment or patient that underwent surgery.[21,22] At the moment, a clinical trail (LIRIC) assesses differences in outcome comparing anti-TNFα treatment with Ileocaecal resection in patients with CD not responsive to prednisolone.[23]

While it has been established that anti-TNFα and/or immunomodulators are generally used to maintain remission in patients with IBD, our study found an association between use of anti-TNFα use and (full) work disability. This could be due to selection bias, as all patients were treated in tertiary referral centers, with most of them having extensive disease involvement (Montreal classification; 47% L3 and 58% E3). Therefore, it is likely that patients receiving anti-TNFα and immunomodulator therapy are patients with more severe disease, receiving “rescue therapy” rather than “top-down therapy”. Smoking was found to be a risk factor for work disability in CD, but not in UC. In this study, smoking status was scored at moment of inclusion, as it has been known that smoking can increase frequency and severity of flares in CD,[24] while it can ameliorate disease in UC.[25] It has been reported that smoking at diagnosis was associated with work disability in UC,[13] whereas others reported on an association between numbers of sick leave days and smoking at the onset of CD,[16] corroborating the current findings.

Clinical risk factor for long-term full work disability

We replicated most known, clinical risk factors for work disability and evaluated risk factors for long-term full work disability (> 80% work disability) comparing it to partial work disability (> 35% work disability). A lower education level was the only factor, which remained statistically significantly associated with long-term full work disability in CD. In only a few studies, educational level has been taken into account showing conflicting results. A higher education level seemed to be associated to work disability and sick leave in patients with CD in previous studies.[8,13] On the other hand, a study by Vester-Andersen *et al* reported no patients with CD with a higher education level that were receiving a disability pension.[13] Hence, the relationship between education level and work disability remains unclear, and more population specific studies are needed before hard conclusions can be drawn.[26,27]

Risk factors identified in this report should be interpreted with caution, as the confidence intervals were quite large and data were derived from third-line referral centres. Furthermore, due to the nature of our data we could not distinguish between work disability solely attributable to IBD, or work disability due to a different cause. However, the median age in our study was 40 years for CD patients and 43 years for UC patients, ages at which age-related comorbidities are generally low. Furthermore, we did not score how physically heavy the patients' jobs were, known to contribute to work disability. Moreover, the reason for work disability was unknown. It could have been that sick leave (< 2 years), was caused by another reason than IBD, for example pregnancy leave or psychological health problems. When comparing our data from the Dutch social security system with the more widely used Work Productivity and Activity Impairment Questionnaire (WPAI), the main difference lies within the objectivity of the data from the Dutch social security system, since it is assessed by a physician as opposed to the more subjective WPAI, which is patient reported. A main disadvantage of the data from the Dutch social security system is that it makes no distinction between work disability due to IBD or due to a different cause. Furthermore, the WPAI assesses activity, not solely work related, on a weekly basis, providing more insight in the disease course.

The strength of our study is the extended disease and patient documentation by the clinician.

In conclusion, in this study we show a higher prevalence of work disability in IBD compared to the general Dutch population. Furthermore, we identified risk factors associated with work disability, with a lower level of education being a risk factor for long-term full work disability in CD, and disease complications being a risk factor for long-term full work disability in UC. In future studies web-based follow-up of Patient-Reported Outcome Measurements (PROMs) including clinical disease activity scores, could be promising tool for detecting a decline in work activity due to disease activity.

References

1. Pariente B, Cosnes J, Danese S, *et al*. Development of the Crohn's disease digestive damage score, the Lémann score. *Inflamm Bowel Dis* 2011;17:1415-1422.
2. Casellas F, Arenas JI, Baudet JS, *et al*. Impairment of health-related quality of life in patients with inflammatory bowel disease: a Spanish multicenter study. *Inflamm Bowel Dis* 2005;11:488-496.
3. Russel MG, Pastoor CJ, Brandon S, *et al*. Validation of the Dutch translation of the Inflammatory Bowel Disease Questionnaire (IBDQ): a health-related quality of life questionnaire in inflammatory bowel disease. *Digestion* 1997;58:282-288.
4. Stark R, König HH, Leidl R. Costs of inflammatory bowel disease in Germany. *Pharmacoeconomics* 2006;24:797-814.
5. Lichtenstein GR, Yan S, Bala M, *et al*. Remission in patients with Crohn's disease is associated with improvement in employment and quality of life and a decrease in hospitalizations and surgeries. *Am J Gastroenterol* 2004;99:91-96.
6. Feagan BG, Bala M, Yan S, *et al*. Unemployment and disability in patients with moderately to severely active Crohn's disease. *J Clin Gastroenterol* 2005;39:390-395.
7. Bernklev T, Jahnsen J, Henriksen M, *et al*. Relationship between sick leave, unemployment, disability, and health-related quality of life in patients with inflammatory bowel disease. *Inflamm Bowel Dis* 2006;12:402-412.
8. Boonen A, Dagnelie PC, Feleus A, *et al*. The impact of inflammatory bowel disease on labor force participation: results of a population sampled case-control study. *Inflamm Bowel Dis* 2002;8:382-389.
9. De Boer AG, Bennebroek Evertsz' F, Stokkers PC, *et al*. Employment status, difficulties at work and quality of life in inflammatory bowel disease patients. *Eur J Gastroenterol Hepatol* 2016;28:1130-1136.
10. van der Valk ME, Mangen MJ, Leenders M, *et al*. Risk factors of work disability in patients with inflammatory bowel disease-a Dutch nationwide web-based survey: work disability in inflammatory bowel disease. *J Crohns Colitis* 2014;8:590-597.
11. Høivik ML, Moum B, Solberg IC, *et al*; IBSEN Group. Work disability in inflammatory bowel disease patients 10 years after disease onset: results from the IBSEN Study. *Gut* 2013;62:368-375.
12. Mandel MD, Bálint A, Lovász BD, *et al*. Work disability and productivity loss in patients with inflammatory bowel diseases in Hungary in the era of biologics. *Eur J Health Econ* 2014;15 Suppl 1:S121-S128.
13. Vester-Andersen MK, Prosberg MV, Vind I, *et al*. Low Risk of Unemployment, Sick Leave, and Work Disability Among Patients with Inflammatory Bowel Disease: A 7-year Follow-up Study of a Danish Inception Cohort. *Inflamm Bowel Dis* 2015;21:2296-2303.
14. Neovius M, Arkema EV, Blomqvist P, *et al*. Patients with ulcerative colitis miss more days of work than the general population, even following colectomy. *Gastroenterology* 2013;144:536-543.
15. Netjes JE, Rijken M. Labor participation among patients with inflammatory bowel disease. *Inflamm Bowel Dis* 2013;19:81-91.
16. Siebert U, Wurm J, Gothe RM, *et al.*; Swiss IBD Cohort Study Group. Predictors of temporary and permanent work disability in patients with inflammatory bowel disease: results of the swiss inflammatory bowel disease cohort study. *Inflamm Bowel Dis* 2013;19:847-855.
17. CBS StatLine - Arbeidsongeschiktheidsuitkering per wet; kenmerken uitkeringsontvanger. Available: http://statline.cbs.nl/Statweb/publication/?DM=SLNLPA=80904NEDD1=0D2=1-2D3=1-5D4=0D5=0D6=4HDR=TSTB=G1, G2,G3,G4,G5VW=T.
18. CBS StatLine - Arbeidsdeelname; kerncijfers. Available: http://statline.cbs.nl/Statweb/publication/?DM=SLNLPA=82309NEDD1=0D2=1-2D3=1,4-6,9D4=0D5=59HDR=G4STB=G1,G2,G3,TVW=T.
19. Data Analysis and Statistical Software | Stata. Available: http://www.stata.com/
20. Olivera P, Spinelli A, Gower-Rousseau C, *et al*. Surgical rates in the era of biological therapy: up, down or unchanged? *Curr Opin Gastroenterol* 2017;33:246-253.

21. Meijs S, Gardenbroek TJ, Sprangers MA, *et al*. Health-related quality of life and disability in patients with ulcerative colitis and proctocolectomy with ileoanal pouch versus treatment with anti-TNF agents. *J Crohns Colitis* 2014;8:686-692.

22. van Gennep S, Sahami S, Buskens CJ, *et al*. Comparison of health-related quality of life and disability in ulcerative colitis patients following restorative proctocolectomy with ileal pouch-anal anastomosis versus anti-tumor necrosis factor therapy. *Eur J Gastroenterol Hepatol* 2017;29:338-344.

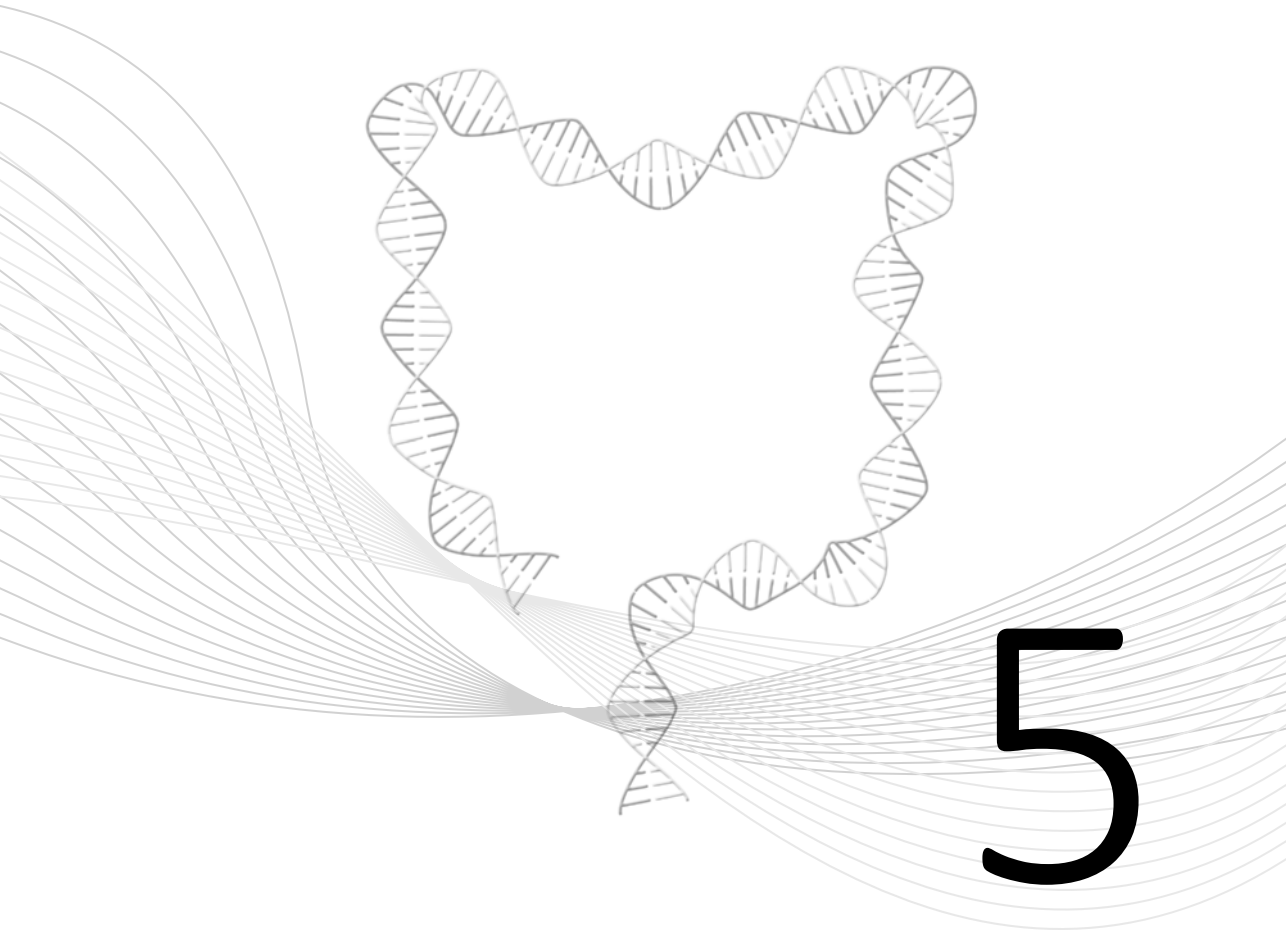
23. LIRIC-trial. Available: https://www.crohn-colitis.nl/wp-content/uploads/2016/09/LIRIC-trial.pdf.

24. Cosnes J. Smoking, physical activity, nutrition and lifestyle: environmental factors and their impact on IBD. *Dig Dis* 2010;28:411-417.

25. Bastida G, Beltrán B. Ulcerative colitis in smokers, non-smokers and ex-smokers. *World J Gastroenterol* 2011;17:2740-2747.

26. Stjernman H, Tysk C, Almer S, *et al*. Unfavourable outcome for women in a study of health-related quality of life, social factors and work disability in Crohn's disease. *Eur J Gastroenterol Hepatol* 2011;23:671-679.

27. Longobardi T, Jacobs P, Bernstein CN. Work losses related to inflammatory bowel disease in the United States: results from the National Health Interview Survey. *Am J Gastroenterol* 2003;98:1064-1072.



5

The impact of ethnicity and country of birth on inflammatory bowel disease phenotype: a prospective cohort study

LM Spekhorst*, M Severs*, NKH de Boer, EAM Festen, HH Fidder, F Hoentjen, F Imhann, DJ de Jong, AE van der Meulen-de Jong, MJ Pierik, CJ van der Woude, G Dijkstra, CY Ponsioen, M Löwenberg, B Oldenburg, RK Weersma;
on behalf of the Parelsnoer Institute (PSI) and the Dutch Initiative on Crohn and Colitis (ICC)
*Authors contributed equally

J Crohns Colitis 2017 Dec 4;11(12):1463-1470

Abstract

Background: The number of patients with inflammatory bowel disease (IBD), of non-Caucasian descent in Western Europe, is increasing. We aimed to explore the impact of ethnicity and country of birth on IBD phenotype.

Methods: IBD patients treated in the eight University Medical Centers in the Netherlands (Dutch IBD Biobank) were divided into two groups according to their ethnicity: 1) Caucasian patients of Western and Central European descent (CEU); and 2) patients of non-Caucasian descent (non-CEU). The non-CEU group was subdivided according to country of birth, into: born in the Netherlands or Western Europe (non-CEU European born); or born outside Western-Europe who migrated to the Netherlands (non-CEU non-European born). Both comparisons were analysed for phenotype differences (by chi-square test).

Results: The Dutch IBD Biobank included 2921 CEU patients and 233 non-CEU patients. Non-CEU Crohn's disease (CD) patients more often had upper gastro-intestinal disease (16% vs 8%, P = 0.001) and anal stenosis (10% vs 4%, P = 0.002) than CEU CD patients. The use of anti-tumour necrosis factor (TNF) agents and immunomodulators was higher in non-CEU IBD patients than in CEU IBD patients (45% vs 38%, P = 0.042) and (77% vs 66%, P = 0.001), respectively. Non-CEU IBD patients born in Europe (*n* = 116) were diagnosed at a lower age than non-CEU IBD patients born outside Europe (*n* = 115) (at 22.7 vs 28.9 years old, P < 0.001).

Conclusion: Non-Caucasians had more severe disease behaviour than Caucasians. Non-CEU patients born in Europe were diagnosed at a lower age with IBD than those born outside Europe who migrated to the Netherlands.

Introduction

The incidence of inflammatory bowel disease (IBD) in Western Europe has remained relatively stable over recent years, but is increasing in developing countries.[1]

It is known that genetic and environmental factors play an important role in the aetiology of IBD. Over recent years, there has been tremendous progress in unravelling the genetic background of IBD, and 200 regions on the human genome have been found to be associated with IBD.[2] Still, the genetic architecture only partly explains the susceptibility to IBD, suggesting that environmental factors contribute to the development of Crohn's disease (CD) and ulcerative colitis (UC).[3] The increasing prevalence of IBD among individuals migrating from low-prevalence regions to high-prevalence regions suggests a role for environmental factors, such as a Westernised diet or lifestyle.[4-6]

Data on ethnic differences regarding phenotypic manifestations of IBD are not consistent, and most published studies so far have relatively small sample sizes.[7-11] Since the number of non-Caucasians (patients from non-European descent) with IBD is increasing in Western European countries,[12] potential phenotypic differences between Caucasian and non-Caucasian IBD patients are of increasing relevance. Treatment paradigms are based on studies predominantly conducted in the Caucasian IBD population,[13] and therefore recognition of ethnic differences in phenotypic manifestations will aid clinical decision making such as optimising individual management and tailoring treatment options. The aim of this study is therefore 2-fold: 1) to gain more insight into the phenotypic differences between Caucasian and non-Caucasian IBD patients in the Netherlands; and 2) to explore the influence of country of birth on the clinical phenotype in non-Caucasian IBD patients by comparing non-Caucasians born in Europe with non-Caucasians born in non-European countries.

Methods

Study design and study population

We conducted a multicentre analysis of IBD patients enrolled in the Dutch IBD Biobank, which is part of the Parelsnoer Institute (www.parelsnoer.org). Since 2007, every IBD patient treated in any one of the eight University Medical Centers (UMCs) is asked to participate in the Dutch IBD Biobank. The Dutch IBD Biobank collects clinical data through an information model that contains 225 IBD-related items. The IBD-related items used for this paper and their definitions can be found in **Supplementary File 1**, available as Supplementary data at *ECCO-JCC* online. These items are retrieved from medical records and are stored in a central database that can be accessed

via web-based trough as a secure transfer. IBD patients are continuously being enrolled in the Dutch IBD Biobank, and data are collected prospectively.[14] At the moment of data freeze (17 July 2014), the Dutch IBD Biobank contained 3388 IBD patients. For this study, patients were excluded if information on ethnicity was missing. Caucasian IBD patients from West or Central European descent were excluded when they were not born in the Netherlands or when their parents were not born in the Netherlands or Europe.

Cohort

The clinician addressed ethnicity/descent during consultation; patients were asked about their descent, country of birth, and country of birth of their parents.

To address the influence of ethnic differences on disease behaviour, IBD patients were divided in two broad groups according to their ethnicity (Figure 1):

1. Caucasians from West or Central European descent (CEU) born in the Netherlands;
2. patients from non-Caucasian descent (non-CEU). Although this second group consists of a heterogeneous group of ethnicities, for the current comparative analysis we will refer to this specific group as “non-CEU”.

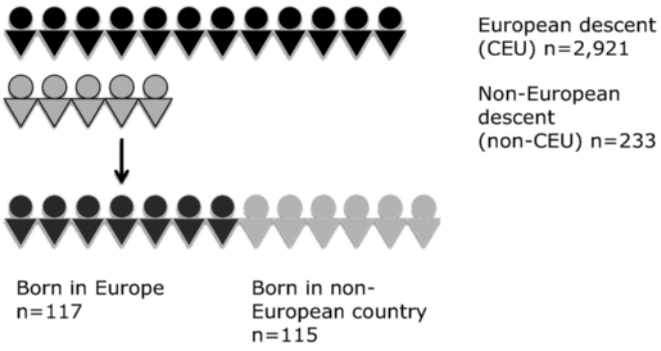


Figure 1 IBD patients were divided according to their ethnicity in two groups CEU and non-CEU. The non-CEU group was divided according to country of birth; born in Europe or born in non-European country.
IBD, inflammatory bowel disease; CEU, Central European descent.

To address the influence of country of birth on phenotypic manifestations of IBD, the second group was then subdivided in two broad groups according to country of birth (Figure 1):

- born in the Netherlands or Western Europe (non-CEU born in Europe);
- born outside Europe and migrated to the Netherlands (non-CEU born in non-European countries).

Statistical analysis

In all analyses, UC patients also included IBD-unclassified (IBD-U) and IBD-indeterminate (IBD-I) patients. Quantitative variables between the different groups were compared using the Mann-Whitney U test. Qualitative variables among the different groups were compared using the Pearson chi-square test or Fisher's exact test. We performed a survival analysis for first IBD-related surgery on time-to-event data using the date of diagnosis and the date of first IBD-related surgery that was registered at the first outpatient clinic visit. Small bowel resection, ileocaecal resection, colon resection, other resections, strictureplasty, ileostomy/colostomy, and surgery for abscesses or fistula were included as first surgery in the analysis. Kaplan-Meier curves for time free of surgery were constructed, and differences between CEU and non-CEU were assessed by the log-rank test. The cumulative proportion of IBD patients remaining free of first IBD-related surgery was calculated at 2 and 5 years after IBD diagnosis. We performed a multivariate Cox proportion hazards analysis for time to first IBD-related surgery for all IBD patients. We corrected for covariates with a P value < 0.05 in the univariate analysis gender (female/male), age, smoking status, ileocolonic disease (L3), extensive colitis (E3), fistulising disease, stricturing disease, penetrating disease, anti-tumour necrosis factor (TNF) use, and immunomodulator use) or when the proportional-hazards assumption was not met. Statistical analyses were performed with Stata Software V.13.1.[15]

Results

Patient population

In total 3154 IBD patients were included in this study (1968 CD, 1108 UC, 72 IBD-U, and 6 IBD-I patients). Of these, 2921 patients were from CEU descent (born in the Netherlands) and 233 patients were from non-CEU descent. Of these 233 patients, 117 patients were from non-CEU descent born in Europe and 115 patients were from non-CEU descent born in non-European countries (for descent and country of birth see **Supplementary File 2**, available as Supplementary data at *ECCO-JCC* online). Parents' countries of origin for the 117 non-CEU descent IBD patients born in Europe are depicted in **Supplementary File 3**, available as Supplementary data at *ECCO-JCC* online.

1. Phenotypic differences between CEU and non-CEU IBD patients

Demographic differences between CEU and non-CEU IBD patients

There were no differences in IBD diagnosis or sex ratio between CEU patients and non-CEU patients. Non-CEU patients were younger at time of inclusion than CEU patients (37.5 vs 42.9 years, P < 0.001).

The median disease duration at inclusion was longer in CEU patients than in non-CEU patients (11.6 vs 10.2 years, P = 0.038) (Table 1).

Age at diagnosis in CEU and non-CEU IBD patients

There was no statistically significant difference in age at diagnosis between CEU patients and non-CEU patients (26.5 vs 25.8 years, P = 0.053) (Table 1).

Table 1 Demographic differences between IBD patients of CEU and non-CEU descent.

	CEU	Non-CEU
n (%)	2921 (100%)	233 (100%)
Sex	2921 (100%)	233 (100%)
Male	1213 (42%)	84 (36%)
Female	1708 (58%)	149 (64%)
Diagnosis	2921 (100%)	233 (100%)
Crohn's disease	1825 (62%)	143 (61%)
UC/IBD-U/IBD-I	1096 (38%)	90 (39%)
Disease characteristics		
Age of inclusion median years (IQR 25-75)	42.9 (32-55)	37.5 (30-49)**
Median Age Diagnosis years (IQR 25-75)	26.5 (20-37)	25.8 (19-33)
Median Disease duration years (IQR 25-75)	11.6 (5-20)	10.2 (5-17)*
Primary sclerosing cholangitis (PSC)[†]	64 (2%)	3 (1.3%)
Family history of IBD[†]	828 (28%)	56 (24%)
Appendectomy[†]	346 (12%)	18 (8%)

[†]Missing values were scored as non-present;
*P < 0.05; **P < 0.001.
CEU: central European descent; UC: ulcerative colitis; IBD-U: inflammatory bowel disease unclassified; IBD-I: inflammatory bowel disease indeterminate; IBD: inflammatory bowel disease; IQR: interquartile range.

Disease behaviour in CEU and non-CEU IBD patients

According to the Montreal classification, division of disease location (L) in CD did not differ between CEU patients and non-CEU patients. Upper gastro-intestinal disease was more common in non-CEU patients than in CEU patients with CD (16% vs 8%, P = 0.001). Disease behaviour in CD was similar between CEU patients and non-CEU patients, except for anal stenosis, which was more common in non-CEU patients (10% vs 4%, P = 0.002). **Supplementary File 4, Table 1** (available as Supplementary data at *ECCO-JCC* online) shows no statistically significant differences in upper gastro-intestinal disease or anal stenosis in a subanalysis between the three largest non-CEU groups (African, Hindustani, Turkish).

No difference was found in disease localisation in UC between CEU patients and non-CEU patients (Table 2).

Medication use in CEU and non-CEU IBD patients

The use of anti-TNFα agents was more frequent in non-CEU IBD patients than in CEU IBD patients (45% vs 38%, P = 0.042). The use of anti-TNFα agents was more frequent in non-CEU patients with UC compared with CEU patients with UC (28% vs 19%, P = 0.047). Immunomodulators were more often used in non-CEU IBD patients than in CEU IBD patients (77% vs 66%, P = 0.001). **Supplementary File 4, Table 1** shows no statistically significant differences in the use of anti-TNFα agents or the use of immunomodulators in a subanalysis between the three largest non-CEU groups (African, Hindustani, Turkish). The use of azathioprine and 6-thioguanine was higher among non-CEU IBD patients than CEU IBD patients (58% vs 49%, P = 0.010) and (7% vs 3%, P = 0.010), respectively. This higher use of azathioprine and 6-thioguanine was predominantly found in non-CEU patients with UC compared with CEU patients with UC (57% vs 41%, P = 0.003) and (11% vs 4%, P = 0.001), respectively (Table 3).

Surgery rates in CEU and non-CEU IBD patients

There was no statistically significant difference in overall surgery rates and the number of stoma and pouches between CEU patients and non-CEU patients (**Supplementary File 4, Table 2**).

At 2 years after IBD diagnosis, the proportion of IBD patients who remained free of IBD surgery was 84% (95% confidence interval (CI): 83%-86%) in CEU patients and 91% (95% CI: 86%-94%) in non-CEU patients (Figure 2). At 5 years after IBD diagnosis, the proportion of IBD patients who remained free of surgery was 74% (95% CI: 72%-76%) in CEU patients and 75% (95% CI: 68%-80%) in non-CEU patients (no statistically significant difference, log-rank test, P = 0.900).

No increased risk of first IBD-related surgery for CEU patients compared with non-CEU patients was found (hazard ratio (HR) 1.01, 95% CI: 0.81-1.27). The multivariate model showed that ileocolonic involvement (L3) (HR 1.48, 95% CI: 1.27-1.72), extensive colitis (E3) (HR 1.18, 95% CI: 1.02-1.38), fistulising disease (HR 1.71, 95% CI: 1.43-2.03), stricturing disease (HR 2.01, 95% CI: 1.72-2.35), penetrating disease (HR 2.35, 95% CI: 1.96-2.81) and anti-TNFα use (HR 1.15, 95% CI: 1.00-1.33) were associated with an increased risk of first IBD-related surgery (**Supplementary File 4, Table 3**).

Table 2 Disease behaviour in IBD patients of CEU and non-CEU descent.

Montreal classification	CEU	Non-CEU
Crohn's disease	1825 (100%)	143 (100%)
A1: diagnosis ≤ 16 years	263 (15%)	28 (20%)
A2: diagnosis 17-40 years	1264 (69%)	95 (66%)
A3: diagnosis > 40 years	298 (16%)	20 (14%)
L1: ileal disease[a]	335 (23%)	27 (25%)
L2: colon disease[a]	451 (31%)	29 (26%)
L3: ileocolon disease[a]	682 (46%)	54 (49%)
L4: upper GI disease[†]	149 (8%)	23 (16%)*
P: peri-anal disease[†]	494 (27%)	35 (24%)
Ulcerative colitis disease		
E1: proctitis[b]	77 (8%)	6 (9%)
E2: left-sided colitis[b]	325 (35%)	21 (31%)
E3: extensive colitis[b]	533 (57%)	41 (60%)
Disease behaviour CD	1825 (100%)	143 (100%)
Fistulising disease		
Enterocutaneous fistula[†]	91 (5%)	8 (6%)
Entero-internal fistula[†]	159 (9%)	12 (8%)
Recto-vaginal fistula[†]	99 (5%)	9 (6%)
Stricturing disease		
Stricture[†]	444 (24%)	28 (20%)
Anal stenosis[†]	75 (4%)	14 (10%)*
Penetrating disease		
Intestinal perforation[†]	52 (3%)	5 (4%)
Intestinal abscess[†]	195 (11%)	17 (12%)

[a]These percentages are calculated for 1468 CEU patients and 110 non-CEU patients with CD;
[b]These percentages are calculated for 935 CEU patients and 68 non-CEU with UC;
[†]Missing values were scored as non-present;
*P < 0.01.
IBD: inflammatory bowel disease; CEU: central European descent; CD: Crohn's disease; GI: gastro-intestinal.

Extraintestinal manifestations in CEU and non-CEU IBD patients

There were no statistically significant differences in the prevalence of skin manifestations, musculoskeletal manifestations, ocular manifestations, osteopenia, or thromboembolic events between CEU patients and non-CEU patients (**Supplementary File 4, Table 4**).

Additional analysis; including all CEU patients independently of country of birth

CEU patients born outside the Netherlands (*n* = 96) were included in the CEU patient group, bringing the total to 3090 CEU IBD patients. Our main findings (i.e., upper gastro-intestinal (GI) disease L4, anal stenosis, anti-TNF use and immunomodulator use) remained statistically significant in these analyses (P < 0.05) comparing CEU and non-CEU patients.

Table 3 Medication use in IBD patients from CEU and non-CEU descent.

	CEU	Non-CEU
Medication use during disease	2861 (100%)	230 (100%)
Anti-TNF[c]	1087 (38%)	103 (45%)*
Anti-TNF CD[a]	880 (49%)	78 (55%)
Anti-TNF UC[b]	207 (19%)	25 (28%)*
Immunomodulators[d]	1898 (66%)	177 (77%)***
Immunomodulators CD[a]	1302 (73%)	114 (81%)*
Immunomodulators UC[b]	596 (56%)	63 (71%)**
Mercaptopurine[e]	463 (16%)	47 (20%)
Mercaptopurine CD[a]	316 (18%)	29 (21%)
Mercaptopurine UC[b]	147 (14%)	18 (20%)
Azathioprine[f]	1413 (49%)	134 (58%)*
Azathioprine CD[a]	971 (54%)	83 (59%)
Azathioprine UC[b]	442 (41%)	51 (57%)**
Thioguanine[g]	94 (3%)	15 (7%)*
Thioguanine CD[a]	53 (3%)	5 (4%)
Thioguanine UC[b]	41 (4%)	10 (11%)***
Methotrexate[h]	365 (13%)	34 (15%)
Methotrexate CD[a]	313 (18%)	28 (20%)
Methotrexate UC[b]	52 (5%)	6 (7%)

[a]These percentages are calculated for 1789 CEU patients and 141 non-CEU patients with CD;
[b]These percentages are calculated for 1072 CEU patients and 89 non-CEU patients with UC;
[c]Anti-TNF: patients used infliximab, adalimumab or certolizumab;
[d]Immunomodulators: patients used mercaptopurine, Puri-nethol, azathioprine, Imuran, thioguanine, methotrexate or Methoject;
[e]Mercaptopurine: patients used mercaptopurine or Puri-nethol;
[f]Azathioprine: patients used azathioprine or Imuran;
[g]Thioguanine: patients used thioguanine or Lanvis;
[h]Methotrexate: patients used methotrexate or Methoject;
*P < 0.05; **P < 0.01; ***P = 0.001.
IBD: inflammatory bowel disease; CEU: central European descent; CD: Crohn's disease; UC: ulcerative colitis; TNF: tumour necrosis factor.

2. Phenotypic differences between non-CEU IBD patients born in Europe and non-CEU IBD patients born in non-European countries who migrated to the Netherlands

Demographic differences between non-CEU IBD patients born in Europe and non-CEU IBD patients born in non-European countries

There were no differences in IBD diagnosis or sex-ratio between non-CEU patients born in Europe and non-CEU patients born in non-European countries. Non-CEU patients born in Europe were younger at time of inclusion than non-CEU patients born in non-European countries (32.0 vs 43.1

years, P < 0.001). The median disease duration at inclusion was longer in non-CEU patients born in non-European countries compared with European countries (11.7 vs 8.0 years, P = 0.006) (Table 4).

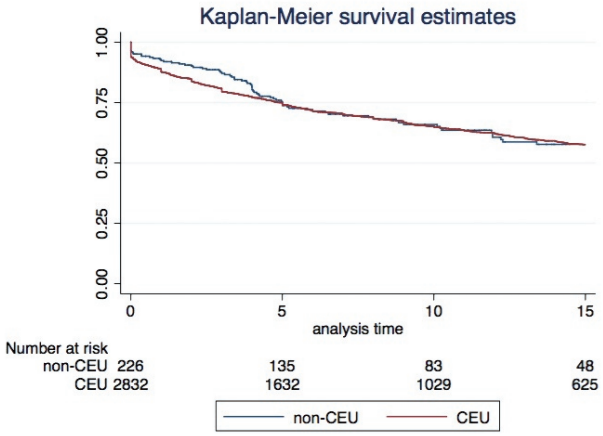


Figure 2 Kaplan-Meier curve; time till first IBD = related surgery compared between CEU and non-CEU patients.

Age at diagnosis in non-CEU IBD patients born in Europe and non-CEU IBD patients born in non-European countries

Non-CEU IBD patients born in Europe were diagnosed at a lower age than non-CEU IBD patients born in non-European countries who migrated to the Netherlands (22.7 vs 28.9 years old, P < 0.001) (Table 4).

Disease behaviour in non-CEU IBD patients born in Europe and non-CEU IBD patients born in non-European countries

Non-CEU patients born in Europe were diagnosed with IBD before the age of 17 years more often compared with non-CEU patients born in non-European countries (29% vs 9%, P = 0.002). Non-CEU patients born in non-European countries were more often diagnosed with IBD after the age of 40 years (22% vs 7%, P = 0.008). According to the Montreal classification, division of disease location (L) in CD did not differ between non-CEU patients born in Europe and non-CEU patients born in non-European countries. No difference was found in disease localisation in UC between non-CEU patients born in Europe and non-CEU patients born in non-European countries (Table 5).

Table 4 Demographic differences between non-CEU IBD patients born in Europe and non-CEU IBD patients born in non-European countries.

	NL/EU born	born non-European country
n (%)	117 (50%)	115 (50%)
Sex	117 (100%)	115 (100%)
Male	43 (37%)	40 (35%)
Female	74 (63%)	75 (65%)
Diagnosis	117 (100%)	115 (100%)
Crohn's disease	75 (64%)	68 (59%)
UC/IBD-U/IBD-I	42 (36%)	47 (41%)
Disease characteristics		
Age of inclusion median years (IQR 25-75)	32.0 (26-44)	43.1 (36-55)**
Median Age Diagnosis years (IQR 25-75)	22.7 (17-29)	28.9 (23-40)**
Median Disease duration years (IQR 25-75)	8.0 (5-14)	11.7 (6-20)*
Primary sclerosing cholangitis (PSC) †	2 (1.7%)	1 (0.9%)
Family history of IBD†	30 (26%)	26 (23%)
Appendectomy†	7 (6%)	11 (10%)

†Missing values were scored as non-present;
*P < 0.01; **P < 0.001.
NL: Netherlands; EU: Europe; CEU: Central European descent; UC: ulcerative colitis; IBD-U: inflammatory bowel disease unclassified; IBD-I: inflammatory bowel disease indeterminate; IBD: inflammatory bowel disease; IQR: interquartile range.

Medication use in non-CEU IBD patients born in Europe and non-CEU IBD patients born in non-European countries

There were no differences in medication use between non-CEU IBD patients born in Europe and non-CEU IBD patients born in non-European countries. Except for the use of mercaptopurine, which was higher among non-CEU patients with UC born in non-European countries than non-CEU patients with UC born in European countries (30% vs 10%, P = 0.015) (Table 6).

Surgery rates in non-CEU IBD patients born in Europe and non-CEU IBD patients born in non-European countries

There was no statistically significant difference in overall surgery rates and the numbers of stoma and pouches between non-CEU IBD patients born in Europe and non-CEU IBD patients born in a non-European country (**Supplementary File 4, Table 5**).

Table 5 Disease behaviour in non-CEU IBD patients born in Europe and non-CEU IBD patients born in non-European countries.

Montreal classification	NL/EU born	born non-European country
Crohn's disease	75 (100%)	68 (100%)
A1: diagnosis ≤ 16 years	22 (29%)	6 (9%)*
A2: diagnosis 17-40 years	48 (64%)	47 (69%)
A3: diagnosis > 40 years	5 (7%)	15 (22%)*
L1: ileal disease[a]	15 (26%)	12 (23%)
L2: colon disease[a]	17 (30%)	12 (23%)
L3: ileocolon disease[a]	25 (44%)	29 (54%)
L4: upper GI disease[†]	11 (15%)	12 (18%)
P: peri-anal disease[†]	16 (21%)	19 (28%)
Ulcerative colitis disease		
E1: proctitis[b]	1 (3%)	4 (12%)
E2: left-sided colitis[b]	9 (26%)	12 (36%)
E3: extensive colitis[b]	24 (71%)	17 (52%)
Disease behaviour CD	75 (100%)	68 (100%)
Fistulising disease		
Enterocutaneous fistula[†]	2 (3%)	6 (9%)
Entero-internal fistula[†]	5 (7%)	7 (10%)
Recto-vaginal fistula[†]	4 (5%)	5 (7%)
Stricturing disease		
Stricture[†]	13 (17%)	15 (22%)
Anal stenosis[†]	10 (13%)	4 (6%)
Penetrating disease		
Intestinal perforation[†]	2 (3%)	3 (4%)
Intestinal abscess[†]	11 (15%)	6 (9%)

[a]These percentages are calculated for 57 non-CEU patients born in NL/EU and 53 non-CEU patients born in non-European countries with CD;
[b]These percentages are calculated for 34 non-CEU patients born in NL/EU and 33 non-CEU patients born in non-European countries with UC;
[†]Missing values were scored as non-present.
*P < 0.01.
NL: Netherlands; EU: Europe; CEU: Central European descent; CD: Crohn's disease. GI: gastro-intestinal.

Extraintestinal manifestations in non-CEU IBD patients born in Europe and non-CEU IBD patients born in non-European countries

There were no statistically significant differences in the prevalence of skin manifestations, musculoskeletal manifestations, ocular manifestations, osteopenia, or thromboembolic events between non-CEU IBD patients born in Europe and non-CEU IBD patients born in a non-European country (**Supplementary File 4, Table 6**).

Table 6 Medication use in non-CEU IBD patients born in Europe and non-CEU IBD patients born in non-European countries.

	NL/EU born	born non-European country
Medication use during disease	115 (100%)	114 (100%)
Anti-TNF[c]	56 (49%)	47 (41%)
Anti-TNF CD[a]	43 (59%)	35 (51%)
Anti-TNF UC[b]	13 (31%)	12 (26%)
Immunomodulators[d]	89 (77%)	87 (76%)
Immunomodulators CD[a]	58 (79%)	56 (82%)
Immunomodulators UC[b]	31 (74%)	31 (67%)
Mercaptopurine[e]	20 (17%)	27 (24%)
Mercaptopurine CD[a]	16 (22%)	13 (19%)
Mercaptopurine UC[b]	4 (10%)	14 (30%)*
Azathioprine[f]	69 (60%)	64 (56%)
Azathioprine CD[a]	44 (60%)	39 (57%)
Azathioprine UC[b]	25 (60%)	25 (54%)
Thioguanine[g]	7 (6%)	8 (7%)
Thioguanine CD[a]	3 (4%)	2 (3%)
Thioguanine UC[b]	4 (10%)	6 (13%)
Methotrexate[h]	17 (15%)	17 (15%)
Methotrexate CD[a]	14 (19%)	14 (21%)
Methotrexate UC[b]	3 (7%)	3 (7%)

[a]These percentages are calculated for 73 non-CEU patients born in NL/EU and 68 non-CEU patients born in non-European countries with CD;
[b]These percentages are calculated for 42 non-CEU patients born in NL/EU and 46 non-CEU patients born in non-European countries with UC;
[c]Anti-TNF: patients used infliximab, adalimumab or certolizumab;
[d]Immunomodulators: patients used mercaptopurine, Puri-nethol, azathioprine, Imuran, thioguanine, methotrexate, or Methoject;
[e]Mercaptopurine: patients used mercaptopurine or Puri-nethol;
[f]Azathioprine: patients used azathioprine or Imuran;
[g]Thioguanine: patients used thioguanine or Lanvis;
[h]Methotrexate: patients used methotrexate or Methoject;
*$P < 0.05$.
NL: Netherlands; EU: Europe; CEU: Central European descent; CD: Crohn's disease; UC: ulcerative colitis.

Discussion

In this large prospective cohort study, we aimed to explore the impact of ethnicity and country of birth on phenotype in IBD patients. We observed that non-CEU patients have more upper GI disease, more stricturing disease and use anti-TNF agents and immunomodulators more often than CEU patients. Furthermore, we found that non-CEU IBD patients born in Europe were diagnosed

at a lower age than non-CEU IBD patients born in non-European countries who migrated to the Netherlands.

Worse disease behaviour in non-CEU compared with CEU IBD patients

We found upper GI disease and anal stenosis to be more common in non-CEU patients compared with CEU patients with CD after a median disease duration of 10 years. Most studies reporting on progression of CD behaviour are conducted in Western populations.[16,17] One study compared disease behaviour between Asian and Australian patients with CD and found stricturing disease to be more common in Asians with CD at diagnosis. Furthermore, they reported a cumulative probability of 20% that CD behaviour would change from inflammatory to stricturing or penetrating disease independent of ethnicity.[9] A review from the UK showed a more severe disease in Asian migrants compared with Caucasian non-migrants.[18] In our cohort, we had six CD patients from Asian descent; one of them had an anal stenosis. However, we were not able to conduct a statistical analysis due to the small sample size. On the other hand, a large review stated that there seems to be no difference in complicated disease behaviour among IBD patients from African American, Asian, and Hispanic descent living in the USA.[19] Interestingly, a study from Belgium compared Moroccan migrants with Caucasian migrants, the latter being a diverse group of Eastern European and Middle Eastern migrants, and found a more severe disease behaviour (more penetrating disease and anti-TNF use) among Moroccan migrants.[20] It has been suggested that a more severe disease behaviour among non-Caucasians is a consequence of delayed disease presentation. However, in our study we could not detect differences in disease behaviour between non-CEU IBD patients born in Europe or non-CEU IBD patients born in non-European countries. A more severe disease behaviour due to a delayed disease presentation is therefore unlikely. To date, genetic risk variants explain only about 10% of the disease risk, which suggests that environmental factors play an important role in the disease pathogenesis. A study from England has shown that disease localisation follows the pattern of the indigenous population after one generation,[21] but other articles challenge these findings.[22] As environmental risk factors play an important role in IBD, a severe disease phenotype could therefore concur with the duration of exposure to these risk factors.

Higher use of IBD medication in non-CEU compared with CEU IBD patients

We found non-CEU IBD patients to use anti-TNF agents and immunomodulators more frequently than CEU patients. In previous studies, the use of medical therapy was similar in Asians compared with Australians.[9,23] Higher anti-TNFα use in our cohort could be explained by the more severe disease behaviour (i.e., more stricturing disease) in non-CEU patients with CD in our cohort. Indeed, a previous study found no differences in infliximab use between African Americans and Caucasian CD patients by only selecting patients with penetrating Crohn's disease.[24] There were no major

differences in extraintestinal manifestations (EIMs) or surgery rates, the latter previously being attributed to differences in health insurance status.[19] However, as inhabitants of the Netherlands all have mandatory health insurance, this possible explanation does not suffice for our cohort. In line with our results no differences in colectomy rates were found between Asians and Caucasians, but interestingly a subgroup analysis across ethnic groups between Indian, Pakistani, Bangladeshi (Asians), and Caucasians did show significant differences in colectomy rates.[25]

Difference in age of onset IBD between non-CEU IBD patients born in Europe compared with non-CEU IBD patients born in non-European countries

Changes in lifestyle, such as a Westernised diet, have been implicated to play a role in the aetiology of IBD.[4-6] In our cohort, we were able to assess the effect of exposure to a Western lifestyle. We found that IBD patients with non-CEU descent born in Europe were diagnosed with IBD at a lower age. This important finding made us speculate that a Western lifestyle might trigger IBD onset more early in life in a genetic susceptible patient; however we were not able to deduce a causal relationship. We realise that the patients in our non-CEU IBD population are genetically very heterogeneous; with most patients having mixed descent (*n* = 72), being African (*n* = 37), or Hindustani (*n* = 37). On the other hand, the non-CEU IBD patients born in non-European countries all have been exposed to a non-Western lifestyle before migrating to the Netherlands. We were not able to address at what age the patient migrated to the Netherlands or whether the diagnosis of IBD was diagnosed in the Netherlands or country of birth. Although phenotype differences were studied in a large cohort with a total of 2921 CEU IBD patients, we had a relatively lower total number of 233 non-CEU IBD patients. It was therefore not possible to conduct a statistical analysis with enough power to detect differences associated with a particular non-CEU population or in a particular non-European country. Still, IBD research in other ethnicities than patients with an European ancestry is of utmost importance and is currently almost non-existent (except for Asia).[26] It has been established that socioeconomic status is associated with IBD risk.[27,28] Unfortunately, we had no data on the socioeconomic status at our disposal.

In conclusion, this study shows that both ethnicity and country of birth are associated with different phenotypes in IBD patients. Furthermore, we found that non-CEU IBD patients born in Europe, exposed to a Westernised lifestyle, have an earlier age of onset compared with non-CEU IBD patients born outside Europe. In clinical practice, clinicians should be aware of distinct phenotype characteristics in the non-CEU IBD populations.

Supplementary Data

Supplementary data are available at *ECCO-JCC* online.

References

1. Molodecky NA, Soon IS, Rabi DM, *et al*. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology* 2012;142:46-54.e42; quiz e30.
2. Liu JZ, van Sommeren S, Huang H, *et al*.; International Multiple Sclerosis Genetics Consortium; International IBD Genetics Consortium. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet* 2015;47:979-86.
3. Brant SR. Promises, delivery, and challenges of inflammatory bowel disease risk gene discovery. *Clin Gastroenterol Hepatol* 2013;11:22-6.
4. Ng SC, Bernstein CN, Vatn MH *et al*.; Epidemiology and Natural History Task Force of the International Organization of Inflammatory Bowel Disease (IOIBD). Geographical variability and environmental risk factors in inflammatory bowel disease. *Gut* 2013;62:630-49.
5. Ananthakrishnan AN. Epidemiology and risk factors for IBD. *Nat Rev Gastroenterol Hepatol* 2015;12:205-17.
6. Ng SC, Tang W, Leong RW, *et al*.; Asia-Pacific Crohn's and Colitis Epidemiology Study ACCESS Group. Environmental risk factors in inflammatory bowel disease: a population-based case-control study in Asia-Pacific. *Gut* 2015;64:1063-71.
7. Basu D, Lopez I, Kulkarni A, *et al*. Impact of race and ethnicity on inflammatory bowel disease. *Am J Gastroenterol* 2005;100:2254-61.
8. Malaty HM, Hou JK, Thirumurthi S. Epidemiology of inflammatory bowel disease among an indigent multi-ethnic population in the United States. *Clin Exp Gastroenterol* 2010;3:165-70.
9. Ng SC, Zeng Z, Niewiadomski O, *et al*. Early course of inflammatory bowel disease in a population-based inception cohort study from 8 countries in Asia and Australia. *Gastroenterology* 2016;150:86-95.e3; quiz e13-4.
10. Nguyen GC, Torres EA, Regueiro M, *et al*. Inflammatory bowel disease characteristics among African Americans, Hispanics, and non-Hispanic Whites: characterization of a large North American cohort. *Am J Gastroenterol* 2006;101:1012-23.
11. Damas OM, Jahann DA, Reznik R, *et al*. Phenotypic manifestations of inflammatory bowel disease differ between Hispanics and non-Hispanic whites: results of a large cohort study. *Am J Gastroenterol* 2013;108:231-9.
12. Centraal Bureau voor de Statistiek. Bevolking; generatie, geslacht, leeftijd en herkomstgroepering, 1 januari. July 18, 2017. http://statline.cbs.nl/statweb/publication/?vw = t&dm = slnl&pa = 37325&d1 = a&d2 = 0&d3 = 0&d4 = 0&d5 = 0-4,137,152,220,237&d6 = 0,4,9,14,18-19&hd = 151214-1201&hdr = g2,g1,g3,t&stb = g4,g5 Accessed February 28, 2017.
13. Hanauer SB, Feagan BG, Lichtenstein GR, *et al*.; ACCENT I Study Group. Maintenance infliximab for Crohn's disease: the ACCENT I randomised trial. *Lancet* 2002;359:1541-9.
14. Parelsnoer Instituut. Referentiekader Parelsnoer Instituut. 2017. http://www.parelsnoer.org/page/nl/Standaardprocedures.
15. Stata. Data Analysis and Statistical Software. http://www.stata.com/
16. Thia KT, Sandborn WJ, Harmsen WS, *et al*. Risk factors associated with progression to intestinal complications of Crohn's disease in a population-based cohort. *Gastroenterology* 2010;139:1147-55.
17. Tarrant KM, Barclay ML, Frampton CM, *et al*. Perianal disease predicts changes in Crohn's disease phenotype-results of a population-based study of inflammatory bowel disease phenotype. *Am J Gastroenterol* 2008;103:3082-93.
18. Misra R, Faiz O, Arebi N. P657 IBD in migrant South Asians: systematic review of epidemiology and disease phenotype. *J Crohns Colitis* 2015;9[Suppl 1]:S410.
19. Afzali A, Cross RK. Racial and ethnic minorities with inflammatory bowel disease in the United States: a systematic review of disease characteristics and differences. *Inflamm Bowel Dis* 2016;22:2023-40.
20. Bouhadan S, Moreels TG. Ethnic differences in inflammatory bowel diseases. *J Gastroint Dig Syst* 2014;4:173.

21. Carr I, Mayberry JF. The effects of migration on ulcerative colitis: a three-year prospective study among Europeans and first- and second- generation South Asians in Leicester [1991–1994]. *Am J Gastroenterol* 1999;94:2918-22.

22. Ko Y, Butcher R, Leong RW. Epidemiological studies of migration and environmental risk factors in the inflammatory bowel diseases. *World J Gastroenterol* 2014;20:1238-47.

23. Ng SC, Tang W, Ching JY, *et al*. Incidence and phenotype of inflammatory bowel disease based on results from the Asia-pacific Crohn's and colitis epidemiology study. *Gastroenterology* 2013;145:158-165.e2.

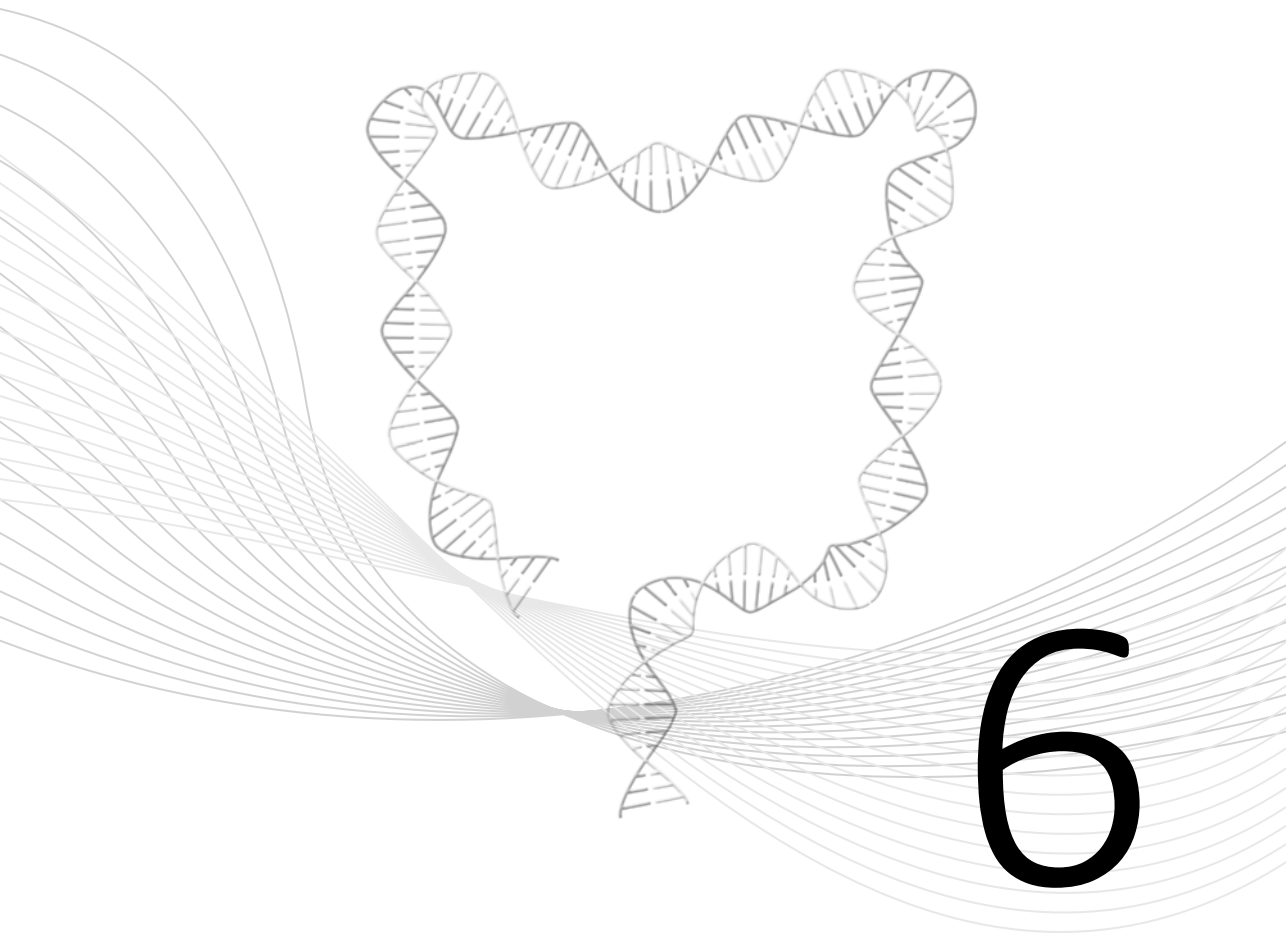
24. Nguyen GC, LaVeist TA, Harris ML, *et al*. Racial disparities in utilization of specialist care and medications in inflammatory bowel disease. *Am J Gastroenterol* 2010;105:2202-8.

25. Misra R, Askari A, Faiz O, *et al*. Colectomy rates for ulcerative colitis differ between ethnic groups: results from a 15-year nationwide cohort study. *Can J Gastroenterol Hepatol* 2016;2016:8723949.

26. Franke A. Inflammatory bowel disease: a global disease that needs a broader ensemble of populations. *Gastroenterology* 2017;152:14-6.

27. Li X, Sundquist J, Hemminki K, *et al*. Risk of inflammatory bowel disease in first- and second-generation immigrants in Sweden: a nationwide follow-up study. *Inflamm Bowel Dis* 2011;17:1784-91.

28. Blanchard JF, Bernstein CN, Wajda A, *et al*. Small-area variations and sociodemographic correlates for the incidence of Crohn's disease and ulcerative colitis. *Am J Epidemiol* 2001;154:328-35.



6

Sex-related differences in patients with inflammatory bowel disease: results of 2 prospective cohort studies

M Severs*, LM Spekhorst*, M-JJ Mangen, G Dijkstra, M Löwenberg, F Hoentjen, AE van der Meulen-de Jong, M Pierik, C Y Ponsioen, G Bouma, CJ van der Woude, ME van der Valk, MJL Romberg-Camps, CHM Clemens, P van de Meeberg, N Mahmmod, JM Jansen, B Jharap, RK Weersma, B Oldenburg, EAM Festen#, and HH Fidder#; on behalf of the Parelsnoer Institute (PSI), the COIN study group and the Dutch Initiative on Crohn and Colitis (ICC)

*#Authors contributed equally

Inflamm Bowel Dis 2018 May 18;24(6):1298-1306

Abstract

Background: The understanding of gender differences in inflammatory bowel disease (IBD) patients is an important step towards tailored treatment for the individual patient. The aim of this study was to compare disease phenotype, clinical manifestations, disease activity, and healthcare utilization between men and women with Crohn's disease (CD) and ulcerative colitis (UC).

Methods: Two multicenter observational cohort studies with a prospective design were used to explore the differences between men and women regarding demographic and phenotypic characteristics and healthcare utilization. Detailed data on IBD-phenotype was mainly available from the Dutch IBD Biobank, while the COIN cohort provided healthcare utilization data.

Results: In the Dutch IBD Biobank study, 2118 CD patients and 1269 UC patients were analysed. Female CD patients were more often current smokers, and male UC patients were more often previous smokers. Early onset CD (< 16 years) was more frequently encountered in males than in females (20% versus 12%, $P < 0.01$). Male CD patients were more often diagnosed with ileal disease (28% versus 20%, $P < 0.01$) and underwent more often small bowel and ileocaecal resection. Extraintestinal manifestations (EIMs) were more often encountered in female IBD patients. In the COIN study, 1139 CD patients and 1213 UC patients were analysed. Male CD patients used prednisone more often and suffered more often from osteopenia. IBD-specific healthcare costs did not differ between male and female IBD patients.

Conclusions: Sex differences in patients with IBD include age of onset, disease location, and EIM prevalence. No large differences in therapeutic management of IBD were observed between men and women with IBD.

Introduction

Inflammatory bowel diseases (IBD), such as Crohn's disease (CD) and ulcerative colitis (UC), are immune-mediated chronic inflammatory diseases of the gastrointestinal tract. More than 1 million residents from the United States of America and 2.5 million Europeans are estimated to suffer from IBD.[1]

Treatment strategies for IBD are aimed at inducing and maintaining long-term remission.[2,3] In other chronic conditions, such as ischemic heart disease, it is recognized that the treatment of men and women requires a different approach.[4] Although incidence rates for male and female IBD patients are well established (approximately 1:1.5 in CD and 1:1.2 in UC, respectively), data on sex-specific differences with respect to disease characteristics of IBD are limited.[5-7] Female sex has been reported as a risk factor for extraintestinal manifestations (EIMs).[5] Studies focusing on predictors for postoperative recurrence of IBD have incorporated sex as a potential predictor, but these studies report conflicting outcomes regarding sex.[8-10] Data regarding differences in healthcare consumption between male and female IBD patients are very limited. One study that compared the use of treatment modalities between hospitalized female and male adolescent CD patients did not observe any significant difference regarding medical procedures, such as endoscopies and blood transfusions.[11] Also in pediatric IBD patients, no differences have been found between males and females regarding the risk for first gastrointestinal surgery.[11-13] The understanding of differences in disease characteristics and current clinical approach between female and male IBD patients is an important step toward tailored treatment in the individual patient.[14,15]

In this study, we aim to compare the phenotype, clinical manifestations, disease course, medical treatment, and other consumption of the healthcare system between adult male and female IBD patients, using 2 large independent multicenter observational cohorts.

Methods

Dutch IBD Biobank

The Dutch IBD Biobank is part of the Parelsnoer Institute[16] and was founded in 2007 with the aim to facilitate basic science and clinical research by providing high-quality biomaterials and an extensive patient data collection. Every adult patient diagnosed with IBD and treated in one of the 8 university medical centers (UMCs) in the Netherlands is eligible for inclusion. Data are collected prospectively by using a standardized information model containing 225 IBD-related items. For the present study, demographic items, diagnosis, smoking status, disease location, disease behaviour,

surgery-related items, medication use, EIMs, and disease complications were used. Definitions of these items can be found in **Supplementary Methods Table 1**.

COIN study

The COIN study (costs of inflammatory bowel disease in the Netherlands) is a prospective multicenter study initiated in 2010, designed to assess costs and health-related quality of life (HrQoL).[17] CD and UC patients who were 18 years or older and attending the IBD units from 7 university medical centers (UMCs) and 7 general hospitals were eligible for participation. The study design has previously been described in detail.[17] In short, data were collected through a web-based baseline questionnaire, followed by quarterly questionnaires. For this study, data regarding demographics, smoking status, disease course and EIMs were collected at baseline. And IBD-specific healthcare consumption (medication use, use of diagnostics, outpatient clinic visits, hospital admissions, and surgical procedures) was measured after 3 months of follow-up. Healthcare costs were calculated by multiplying units of self-reported healthcare utilization during follow-up by their corresponding unit prices.[18] Total healthcare costs consisted of medication use, hospital admissions, surgeries, diagnostic procedures, and outpatient clinic visits. Costs were expressed in Euros for the year 2015. Definitions of all above-mentioned items can be found in **Supplementary Methods Table 2**.

Analysing the Dutch IBD Biobank and COIN cohort

To prevent duplicates, only patients in the COIN cohort that did not participate in the Dutch IBD Biobank cohort were included in the analysis. The 2 cohorts have a different study aim. The Dutch IBD Biobank was founded to facilitate basic science and clinical research by providing high-quality biomaterials and an extensive patient data collection. The Dutch IBD Biobank is rigorous and detailed in phenotype collection (225 IBD items). The COIN study was designed to assess costs and health-related quality of life (HrQoL). Furthermore, the cohorts differ regarding patient inclusion, data collection, and selection of patients. In the Dutch IBD Biobank, patients were included and followed during hospital visits for IBD. Data in the COIN were collected through a web-based baseline questionnaire, followed by quarterly questionnaires. Patients were invited to participate at home. The Dutch IBD Biobank consisted entirely of IBD patients treated in tertiary referral centers (i.e., university hospitals). Patients were included during a hospital visit for IBD. Data collection is based on electronic patient databases. Patients of the COIN study were self-reported through questionnaires and recruited from university hospitals and general hospitals. Because of the distinct features in each of the cohorts, analyses of the Dutch IBD Biobank study and the COIN study were performed separately.

Statistical analysis

In all analyses, IBD-unclassified (IBD-U) and IBD-indeterminate (IBD-I) patients were included in the UC category. All analyses were repeated for CD and UC patients separately. Descriptive statistics were performed to describe differences between the male and female sex groups. Categorical variables were compared by a Pearson $\chi 2$ test, and continuous variables by a Mann-Whitney-U test. Statistical analyses were performed with Stata Software V.13.1 (College Station, TX: Stata Corp LP) and SPSS version 21.0 (Armonk, NY: IBM Corp).

Ethical considerations

The Dutch IBD Biobank was carried out with the approval of both a Central Medical Ethics Committee (MEC) and approval of local MECs of all participating centers. The COIN study was carried out with the approval of the MEC of the University Medical Center, Utrecht.

Results

Patient population and male:female ratio

In the Dutch IBD Biobank, 2118 CD and 1269 UC patients were included. The male:female ratio was 1:1.7 in CD and 1:1.1 in UC (Table 1). In the COIN study, 1139 CD patients and 1213 UC patients were included. The male:female ratio in this study was 1:1.6 in CD and 1:1.1 in UC (Table 1).

Demographic differences between female and male CD and UC patients

In the Dutch IBD Biobank, female CD patients were more often current smokers than male CD patients (43% versus 33%, $P < 0.01$). Male UC patients were more often previous smokers than female UC patients (51% versus 40%, $P < 0.01$) (Table 1). Comparable outcomes were observed in the COIN study (Table 1).

Differences in disease behaviour between female and male CD and UC patients

In the Dutch IBD Biobank, men were diagnosed with CD with early onset (younger than 16 years old) more frequently than women with CD (20% versus 12%, $P < 0.01$) (Table 2). Also, UC was diagnosed more often in men older than 40 years in comparison with women (and 33% versus 22%, $P < 0.01$, respectively). Ileal disease (Montreal L1) was encountered more frequently in male CD patients, while colonic disease (Montreal L2) was more common in female CD patients. Proctitis (Montreal E1) was more frequent in female than in male UC patients. In the COIN study, men with CD or UC were diagnosed at 40 years or older more often than women (29% versus 17% in CD; 42% versus

26% in UC, both P < 0.01) (**Supplementary Table S1**). No differences in the incidence of flares were observed between males and females in either cohort.

Table 1 Demographic differences between female and male IBD patients

Dutch IBD Biobank	CD (n = 2118)			UC (n = 1269)		
	Male	Female	P value	Male	Female	P value
Age at inclusion median years (IQR)	42 (31-56)	41 (31-52)	0.13	49 (36-60)	42 (33-53)	< 0.01
Employed at baseline n (%)	415 (73)	557 (58)	< 0.01	335 (79)	334 (73)	< 0.05
Low education n (%)	366 (48)	753 (58)	< 0.01	297 (51)	335 (52)	0.77
Smoking status n (%)						
Current	138 (33)	351 (43)	< 0.01	62 (21)	68 (17)	0.26
Never	283 (42)	471 (39)	0.23	235 (44)	321 (53)	< 0.01
Previously	259 (48)	392 (45)	0.39	241 (51)	212 (40)	< 0.01
Family history of IBD† n (%)	231 (30)	382 (28)	0.47	154 (26)	165 (25)	0.78
COIN Study	CD (n = 1139)			UC (n = 1213)		
	Male	Female	P value	Male	Female	P value
Age at inclusion mean years (SD)	51 (14)	45 (13)	< 0.01	54 (13)	46 (13)	< 0.01
Employed at baseline n (%)	230 (72)	322 (64)	0.02	347 (80)	337 (81)	0.65
Low education n (%)	266 (60)	479 (69)	< 0.01	371 (59)	363 (62)	0.36
Smoking status n (%)			< 0.01			< 0.01
Current	72 (16)	180 (26)	< 0.01	57 (9)	78 (13)	0.02
Never	233 (53)	331 (48)	0.10	334 (53)	339 (58)	0.12
Previously	138 (31)	185 (27)	0.10	235 (38)	170 (29)	< 0.01
Family history of IBD n (%)	98 (22)	149 (21)	0.95	116 (19)	126 (22)	0.44

†missing values were scored as non-present.
N: number; %: percentage of total; UC: ulcerative colitis; CD: Crohn's disease; IBD: Inflammatory Bowel Disease; SD: standard deviation; IQR: interquartile range.

Extraintestinal manifestations in female and male CD and UC patients

In the Dutch IBD Biobank, skin- and joint-manifestations were more common in female CD and UC patients than in male CD and UC patients (Table 3). Osteopenia and osteoporosis were diagnosed more often in male than in female CD patients (28% versus 21% in females, P < 0.01). After correcting for smoking behaviour, age at diagnosis, and disease duration, female sex was associated with EIMs with an adjusted odds ratio (OR) of 2.3 (95% CI: 1.9–2.8) in CD patients and an adjusted OR of 1.5 (95% CI: 1.1–2.3) in UC patients (Supplementary Table S2). In the COIN study,[19] a similar trend was observed, although it did not reach statistical significance (**Supplementary Table S3**).

Surgery rates in female and male CD and UC patients

In the Dutch IBD Biobank, more male CD patients underwent small bowel and ileocaecal resection than female CD patients (16% versus 9%, and 40% versus 33%, respectively, both P < 0.01) (Table

4). No further differences in surgery rates between men and women (colon resection, ileostomy, colostomy, abscess/fistula surgery, and pouches) were observed. These observations were confirmed in the COIN study (**Supplementary Table S4**).

Table 2 Disease behaviour in female and male IBD patients in the Dutch IBD Biobank

	CD (n = 2118)			UC (n = 1269)		
N (%)	Male	Female	P value	Male	Female	P value
Type of IBD			< 0.01			< 0.01
CD	773 (37)	1345 (63)				
UC				604 (48)	665 (52)	
Disease duration at baseline median, years (IQR)	13 (6-22)	12 (6-21)	0.16	11 (5-20)	10 (5-18)	0.19
Montreal classification						
A1: diagnosis ≤ 16 years	153 (20)	166 (12)	< 0.01	54 (9)	80 (12)	0.07
A2: diagnosis 17-40 years	487 (63)	980 (73)	< 0.01	348 (58)	437 (66)	< 0.01
A3: diagnosis > 40 years	133 (17)	199 (15)	0.14	202 (33)	148 (22)	< 0.01
L1: ileal disease[a]	169 (28)	210 (20)	< 0.01			
L2: colon disease[a]	162 (26)	356 (33)	< 0.01			
L3: ileocolon disease[a]	280 (46)	500 (47)	0.67			
L4: upper GI disease[†]	76 (10)	101 (8)	0.06			
P: peri-anal disease[†]	221 (29)	342 (25)	0.11			
E1: proctitis[b]				26 (5)	59 (10)	< 0.01
E2: left-sided colitis[b]				179 (36)	197 (35)	0.86
E3: extensive colitis[b]				299 (59)	307 (55)	0.11
Disease behaviour CD						
Fistulising[†]	102 (13)	236 (18)	< 0.01			
Stricturing[†]	216 (28)	341 (25)	0.19			
Penetrating[†]	110 (14)	159 (12)	0.11			

[a]These percentages are calculated for 1677 CD patients (1066 female CD patients and 611 male CD patients);
[b]These percentages are calculated for 1067 UC patients (563 female UC patients and 504 male UC patients);
[†]missing values were scored as non-present.
N: number; %: percentage of total; UC: ulcerative colitis; CD: Crohn's disease; IBD: Inflammatory Bowel Disease; IQR: interquartile range.

Table 3 Extraintestinal manifestations in the Dutch IBD Biobank

	CD (n = 2118)			UC (n = 1269)		
N (%)	Male	Female	P value	Male	Female	P value
Skin manifestations[†]	47 (6)	203 (15)	< 0.01	31 (5)	55 (8)	0.03
Arthropathy[†]	95 (12)	272 (20)	< 0.01	55 (9)	95 (14	< 0.01
Arthritis[†]	27 (3)	119 (9)	< 0.01	25 (4)	43 (6)	0.07
Ocular manifestations[†]	31 (4)	73 (5)	0.15	16 (3)	27 (4)	0.17
Osteopenia[†]	219 (28)	277 (21)	< 0.01	93 (15)	87 (13)	0.24

[†]missing values were scored as non-present.
N: number; %: percentage of total; UC: ulcerative colitis; CD: Crohn's disease.

Table 4 Surgery rates in female and male IBD patients in the Dutch IBD Biobank

	CD (n = 2118)			UC (n = 1269)		
N (%)	Male	Female	P value	Male	Female	P value
Small bowel resection†	127 (16)	115 (9)	< 0.01			
Ileocaecal resection†	308 (40)	450 (33)	< 0.01			
Colon resection†	124 (16)	244 (18)	0.22	102 (17)	121 (18)	0.54
Strictureplasty†	44 (6)	45 (3)	0.01			
Ileostomy/colostomy†	90 (12)	193 (14)	0.08	56 (9)	75 (11)	0.24
Surgery for abscesses or fistula†	160 (21)	307 (23)	0.26	13 (2)	14 (2)	0.95
Stoma†	90 (12)	180 (13)	0.25	63 (10)	69 (10)	0.98
Pouch†	13 (1.7)	25 (1.9)	0.77	52 (9)	65 (10)	0.47
Post operative stricture†	47 (6)	60 (4)	0.10			

†missing values were scored as non-present.
N: number; %: percentage of total; UC: ulcerative colitis; CD: Crohn's disease.

Table 5 Medication use and other healthcare use in female and male IBD patients in the COIN study

	CD (n = 940)*			UC (n = 1023)*		
N (%)	Male	Female	P value	Male	Female	P value
Anti-TNF	81 (22)	118 (21)	0.63	18 (3)	21 (4)	0.37
Adalimumab	35 (10)	67 (12)	0.28	8 (2)	5 (1)	0.54
Infliximab	47 (13)	51 (9)	0.06	10 (2)	16 (3)	0.13
Azathioprine	98 (27)	126 (22)	0.11	80 (15)	58 (12)	0.22
Mercaptoprine	25 (7)	46 (8)	0.47	35 (6)	27 (6)	0.59
Methotrexate	10 (3)	12 (2)	0.54	2 (0)	3 (1)	0.56
Budesonide	16 (4)	31 (5)	0.46	17 (3)	9 (2)	0.20
Prednisone	25 (7)	21 (4)	0.03	32 (6)	24 (5)	0.54
Sulfazalazine	6 (2)	8 (1)	0.79	3 (1)	6 (1)	0.23
Mesalazine	92 (25)	134 (23)	0.60	329 (61)	283 (59)	0.62
Use of diagnostics	89 (24)	123 (22)	0.36	92 (17)	73 (15)	0.45
Hospitalization due to IBD	19 (5)	17 (3)	0.09	14 (3)	18 (4)	0.28
Outpatient clinic visit due to IBD	171 (77)	281 (82)	0.11	234 (82)	197 (77)	0.12

N: number; %: percentage of total; UC: ulcerative colitis; CD: Crohn's disease; *: At t = 3 months of follow-up

Medication use, healthcare consumption, and healthcare costs in female and male CD and UC patients

In the Dutch IBD Biobank, no differences regarding the use of IBD-specific medication (anti-TNFα compounds and immunomodulators) were observed between men and women (**Supplementary Table S5**).

In the COIN study, prednisone use was reported more frequently in male CD patients than in female CD patients (6.8% versus 3.7%, P = 0.03) (Table 5). No other differences regarding the use of IBD-specific medication, diagnostics, the number of hospitalizations, or the number of outpatient

clinic visits between men and women with CD or UC were observed. Quarterly healthcare costs did not differ between male and female CD and UC patients (Table 6).

Table 6 Mean healthcare costs in female and male patients in the COIN, calculated in euros for the year 2015, calculated over 3 months of follow-up

Mean € (95% CI)	Male	Female	P value
Crohn's disease (369 males and 572 females)			
Total healthcare costs	1768 (1502 – 2075)	1520 (1298 – 1746)	0.19
Medication costs	1205(1006– 1411)	1118 (953 – 1286)	0.52
Hospitalization costs	370 (224 – 540)	198 (107 – 300)	0.08
Surgery costs	3 (0 – 11)	14 (3 – 29)	0.27
Diagnostics costs	48 (35 – 62)	46 (34 – 58)	0.89
Outpatient clinic costs	125 (100 – 153)	127 (100 – 165)	0.94
Ulcerative colitis (544 males and 480 females)			
Total healthcare costs	538 (429 – 646)	608 (480 – 739)	0.45
Medication costs	299 (234 – 365)	353 (269 – 447)	0.37
Hospitalization costs	113 (50 – 183)	122 (50 – 195)	0.87
Surgery costs	5 (0 – 12)	12 (0 – 24)	0.37
Diagnostics costs	32 (24 – 42)	36 (25 – 47)	0.61
Outpatient clinic costs	82 (71 – 95)	75 (62 – 88)	0.39

CI: confidence interval

Additional analysis: including all patients in the COIN cohort without removing duplicates

By analysing all variables described above in the COIN study without excluding patients who were also included in the Biobank study, no statistical differences were observed regarding the outcomes (data available on request).

Discussion

In this large nationwide multicenter study based on 2 IBD cohorts, we aimed to identify clinical relevant differences in disease characteristics and current clinical approach between female and male IBD patients. We observed that early onset CD (< 16 years old) was more prevalent in male patients. Ileal disease and small bowel surgery were more common in male than in female CD patients. Both female CD and UC patients suffered more often from EIMs than male patients. Prednisone use was higher in male CD patients compared to female CD patients. The use of other IBD-specific medication, outpatient clinic visits, diagnostic procedures, and hospitalizations did not differ between female and male IBD patients.

Although correction for gender is frequently performed in the analysis of clinical parameters in IBD studies, results of previous studies usually did not correct for sex.[20-22] Our finding in the Dutch IBD Biobank that male sex was associated with a CD diagnosis before the age of 16 years seems to have some base in previous literature.[23] However, this observation was not confirmed in the COIN cohort. Differences in exposure to hormonal changes during pubertal development could underlie the sex-related differences of early onset CD,[24] but environmental differences (such as diet and smoking behaviour) between males and females in puberty may also contribute.[25]

In the Dutch IBD Biobank, the ileum was more frequently affected in men with CD than in women. Prior data on this subject are scarce, with only 1 other study confirming our observation (i.e., ileitis terminal) in 33% of males versus 29% of females with CD.[10] It is likely that the increased prevalence of ileal involvement translated into more small bowel and ileocaecal resections in men with CD. One could argue that higher resection rates reflect more severe disease behaviour in men. However, no other arguments for a more severe course in men were found in the Dutch IBD Biobank, since comparable numbers of male and female patients used anti-TNFα compounds and immunomodulators. Therefore, our data suggests that gender may be a potential risk factor for ileal disease involvement and subsequent ileal surgery. Prior reports regarding the role of gender in relation to (overall abdominal) surgery are conflicting.[13,26]

A multivariable analysis for EIMs in the COIN study has previously been published.[19] In both cohorts, an association was found between female sex and EIMs in IBD patients.[27-30] Smoking is strongly associated with EIMs,[19] and one could hypothesize that the female association with EIMs is explained by the fact that women with IBD were more often smokers than men with IBD.[31] However, multivariable analysis suggests that female sex is an independent risk factor for EIMs, in both CD and UC patients. Similarly, many rheumatic diseases, including rheumatoid arthritis (RA) are more common in women than in men.[32] Sex hormones are thought to underlie this gender difference. In RA, disease activity is found to be correlated with prolactin plasma levels.[33] Prolactin levels have been found to rise in women with RA post-partum, corresponding with a higher incidence of flares.[34] Other suggested mechanisms include a lower level of serum androgens resulting in an increased production of interleukin-2 and suppression of the synthesis of autoantibodies.[35,36] The same mechanisms might explain the higher incidence of EIMs in female IBD patients.

In the COIN study, prednisone was more often prescribed to male CD patients, but since the numbers were small, results need to be interpreted with caution. In regard to the use of other IBD-specific medication, diagnostic procedures, outpatient clinic visits, hospitalizations, and associated costs, we found no gender differences in the COIN study. Only very few previous studies have reported on differences in IBD-related healthcare utilization in men and women. One study did not find a significant difference with respect to procedures (endoscopy, surgery, and blood transfusions) and the use of corticosteroids or biologic agents by comparing treatment between hospitalized

male and female adolescent CD patients.[11] Another study found that female IBD patients were treated with immunosuppressive agents less frequently than male IBD patients.[31] The authors hypothesize that this difference results from the tendency to prescribe less medication to women of child-bearing potential. However, a sub-analysis did not substantiate this assumption. Apart from more ileal and ileocolonic surgery in male CD patients in the Dutch IBD Biobank cohort, no differences in the management of male and female IBD patients were detected. Interestingly, in other immune-mediated diseases, female patients have been postulated to be less responsive to anti-TNFα therapy.[37-39] The current study was not designed to assess response to therapy, therefore we cannot confirm nor refute this finding.

Strengths of this study include the large sample size (n > 5700) and the corroboration of two independent cohorts. Most of the observed differences between men and women were encountered in both cohorts or showed a similar trend, which supports the validity of both the separate cohorts and our conclusions of this current study. Moreover, the same results were obtained when repeating analyses for the complete COIN cohort, (i.e., including the duplicates from the IBD Biobank cohort). Specific strengths of the Dutch IBD Biobank are the standardized and concise method of phenotyping of the included patients. Strengths of the COIN study include the diversity of the case mix and the comprehensive and prospective way consumption of healthcare and associated healthcare costs were studied. Some differences between the 2 cohorts warrant comment. First, the Dutch IBD Biobank study exclusively recruited patients from university medical centers, whereas the COIN study included patients from both academic medical centers and general hospitals. Data collection in the Dutch IBD Biobank was predominantly conducted from medical records, while the data in the COIN study was self-reported. Self-reported information is generally less objective, and this could be considered a limitation. However, self-report is a validated method to study healthcare consumption and incurred healthcare costs in IBD patients.[40] In addition, a non-responder analysis showed no relevant statistical significant differences regarding demographic data (including gender) and disease course items between self-reported responders and non-responders.[17] Therefore, it was concluded that self-report reliably reflects consumption of health care in patients with IBD. With regard to disease characteristics, our findings were primarily based on the rigorous phenotyping from the Dutch IBD Biobank study.

In summary, our study revealed several sex differences between male and female IBD patients regarding age of onset, smoking behaviour, disease location, and EIMs. For example, both cohorts showed that smokers were most often female patients with CD. As smoking is a clinical parameter associated with severe disease course, clinicians should discuss this with every patient. This is an example where our results could aid toward the development of tailored treatment for the individual IBD patient.

Supplementary data

Supplementary data are available at *Inflammatory Bowel Diseases* online.

References

1. Kaplan GG. The global burden of IBD: from 2015 to 2025. *Nat Rev Gastroenterol Hepatol* 2015;12:720-7.
2. Drug Therapy for Crohn's. http://www.gastro.org/guidelines/drug-therapy-for-crohn-s
3. Ulcerative Colitis. http://campaigns.gastro.org/algorithms/UlcerativeColitis/
4. Khamis RY, Ammari T, Mikhail GW. Gender differences in coronary heart disease. *Heart* 2016;102:1142-9.
5. Wagtmans MJ, Verspaget HW, Lamers CB, *et al.* Gender-related differences in the clinical course of Crohn's disease. *Am J Gastroenterol* 2001;96:1541-6.
6. Ozin Y, Kilic MZY, Nadir I, *et al.* Clinical features of ulcerative colitis and Crohn's disease in Turkey. *J Gastrointestin Liver Dis* 2009;18:157-62.
7. Morris T, Rhodes J. Incidence of ulcerative colitis in the Cardiff region 1968-1977. *Gut* 1984;25:846-8.
8. Bernell O, Lapidus A, Hellers G. Risk factors for surgery and postoperative recurrence in Crohn's disease. *Ann Surg* 2000;231:38-45.
9. Hofer B, Böttger T, Hernandez-Richter T, *et al.* The impact of clinical types of disease manifestation on the risk of early postoperative recurrence in Crohn's disease. *Hepatogastroenterology*;48:152-5.
10. Romberg-Camps MJL, Dagnelie PC, Kester ADM, *et al.* Influence of phenotype at diagnosis and of other potential prognostic factors on the course of inflammatory bowel disease. *Am J Gastroenterol* 2009;104:371-83.
11. Dotson JL, Bricker JB, Kappelman MD, *et al.* Assessment of Sex Differences for Treatment, Procedures, Complications, and Associated Conditions Among Adolescents Hospitalized with Crohn's Disease. *Inflamm Bowel Dis* 2015;21:2619-24.
12. Schaefer ME, Machan JT, Kawatu D, *et al.* Factors That Determine Risk for Surgery in Pediatric Patients With Crohn's Disease. *Clin Gastroenterol Hepatol* 2010;8:789-794.e2.
13. Gupta N, Cohen SA, Bostrom AG, *et al.* Risk Factors for Initial Surgery in Pediatric Patients With Crohn's Disease. *Gastroenterology* 2006;130:1069-77.
14. Boyapati RK, Kalla R, Satsangi J, *et al.* Biomarkers in Search of Precision Medicine in IBD. *Am J Gastroenterol* 2016;111:1682-90.
15. Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med* 2015;372:793-5.
16. Parelsnoer Institute - Home. http://en.parelsnoer.org/
17. van der Valk ME, Mangen M-JJ, Leenders M, *et al.* Healthcare costs of inflammatory bowel disease have shifted from hospitalisation and surgery towards anti-TNFα therapy: results from the COIN study. *Gut* 2014;63:72-9.
18. Hakkaart-van Roijen L, Tan SS, Bouwmans CAM. Handleiding voor kostenonderzoek, methoden en standaard kostprijzen voor economische evaluaties in de gezondheidszorg. College voor zorgverzekeringen Geactualiseerde versie. 2010.
19. Severs M, van Erp SJH, van der Valk ME, *et al.* Smoking is Associated With Extra-intestinal Manifestations in Inflammatory Bowel Disease. *J Crohns Colitis* 2016;10:455-61.
20. Solberg IC, Lygren I, Jahnsen J, *et al.* Clinical course during the first 10 years of ulcerative colitis: results from a population-based inception cohort (IBSEN Study). *Scand J Gastroenterol* 2009;44:431-40.
21. Vegh Z, Burisch J, Pedersen N, *et al.* Treatment Steps, Surgery, and Hospitalization Rates During the First Year of Follow-up in Patients with Inflammatory Bowel Diseases from the 2011 ECCO-Epicom Inception Cohort. *J Crohn's Colitis* 2015;9:747-53.
22. Sjöberg D, Holmström T, Larsson M, *et al.* Incidence and clinical course of Crohn's disease during the first year - Results from the IBD Cohort of the Uppsala Region (ICURE) of Sweden 2005-2009. *J Crohn's Colitis* 2014;8:215-22.
23. Benchimol EI, Fortinsky KJ, Gozdyra P, *et al.* Epidemiology of pediatric inflammatory bowel disease: a systematic review of international trends. *Inflamm Bowel Dis* 2011;17:423-39.

24. Herzog D, Buehr P, Koller R, *et al*. Gender Differences in Paediatric Patients of the Swiss Inflammatory Bowel Disease Cohort Study. *Pediatr Gastroenterol Hepatol Nutr* 2014;17:147.

25. Ngo ST, Steyn FJ, McCombe PA. Gender differences in autoimmune disease. *Front Neuroendocrinol* 2014;35:347-69.

26. Basilisco G, Campanini M, Cesana B, *et al*. Risk factors for first operation in Crohn's disease. *Am J Gastroenterol* 1989;84:749-52.

27. Barreiro-de Acosta M, Domínguez-Muñoz JE, Núñez-Pardo de Vera MC, *et al*. Relationship between clinical features of Crohn's disease and the risk of developing extraintestinal manifestations. *Eur J Gastroenterol Hepatol* 2007;19:73-8.

28. Ditisheim S, Fournier N, Juillerat P, *et al*. Inflammatory Articular Disease in Patients with Inflammatory Bowel Disease: Result of the Swiss IBD Cohort Study. *Inflamm Bowel Dis* 2015;21:2598-604.

29. Li Y-C, Li W-Z, Wu C-R, *et al*. Prevalence and characteristics of ophthalmological extra-intestinal manifestations in Chinese patients with inflammatory bowel disease. *Int J Ophthalmol* 2016;9:1476-9.

30. Karmiris K, Avgerinos A, Tavernaraki A, *et al*. Prevalence and Characteristics of Extra-intestinal Manifestations in a Large Cohort of Greek Patients with Inflammatory Bowel Disease. *J Crohn's Colitis* 2016;10:429-36.

31. Blumenstein I, Herrmann E, Filmann N, *et al*. Female patients suffering from inflammatory bowel diseases are treated less frequently with immunosuppressive medication and have a higher disease activity: a subgroup analysis of a large multi-centre, prospective, internet-based study. *J Crohns Colitis* 2011;5:203-10.

32. KVIEN TK, Uhlig T, Ødegård S, *et al*. Epidemiological Aspects of Rheumatoid Arthritis: The Sex Ratio. *Ann N Y Acad Sci* 2006;1069:212-22.

33. Zoli A, Lizzio MM, Ferlisi EM, *et al*. ACTH, cortisol and prolactin in active rheumatoid arthritis. *Clin Rheumatol* 2002;21:289-93.

34. Olsen NJ, Kovacs WJ. Hormones, pregnancy, and rheumatoid arthritis. *J Gend Specif Med*;5:28-37.

35. Barragán-Martínez C, Amaya-Amaya J, Pineda-Tamayo R, *et al*. Gender Differences in Latin-American Patients With Rheumatoid Arthritis. *Gend Med* 2012;9:490-510.e5.

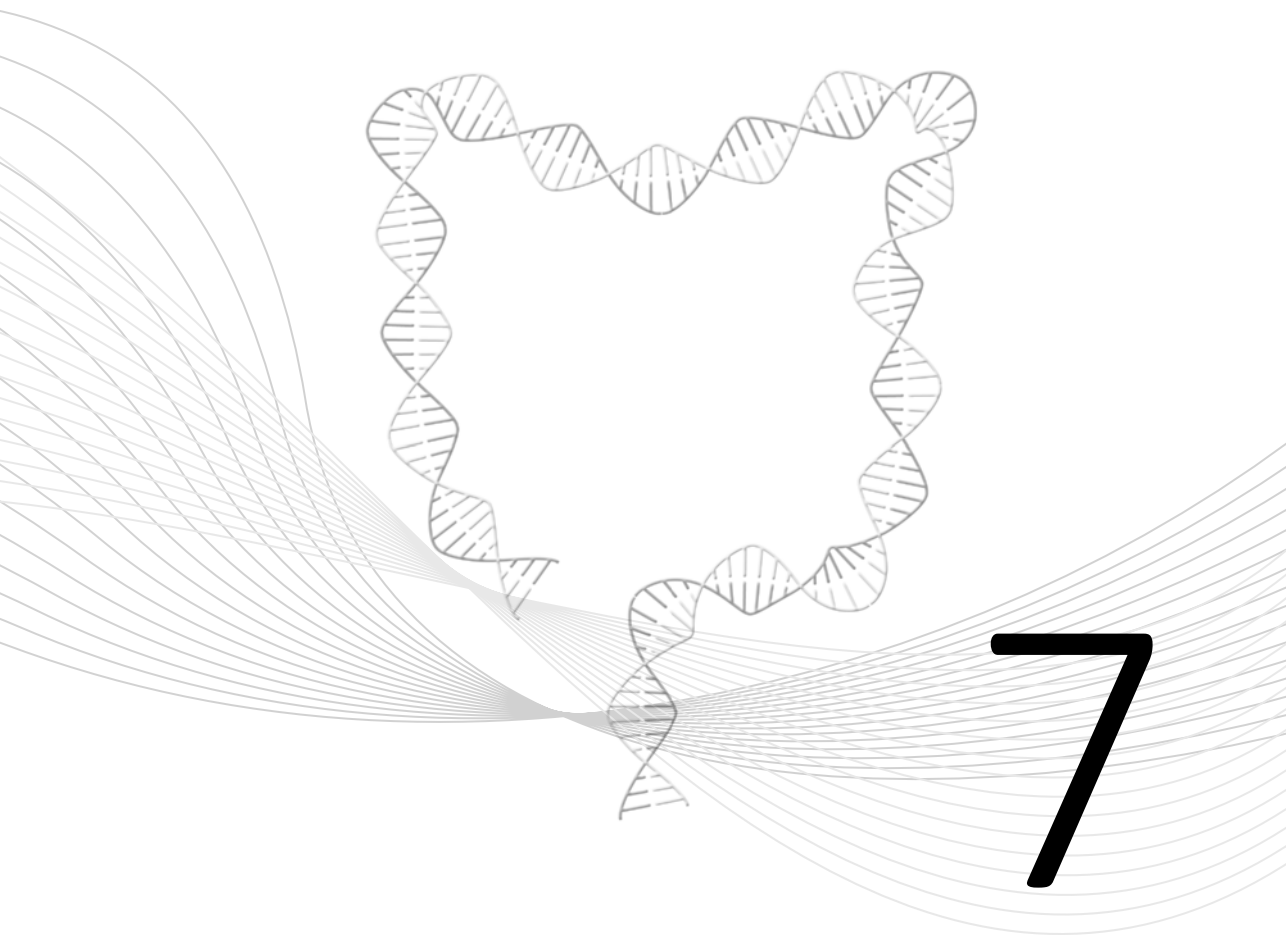
36. Cutolo M, Villaggio B, Craviotto C, *et al*. Sex hormones and rheumatoid arthritis. *Autoimmun Rev* 2002;1:284-9.

37. De Simone C, Caldarola G, Maiorino A, *et al*. Clinical predictors of nonresponse to anti-TNFα agents in psoriatic patients: A retrospective study. *Dermatol Ther* 2016;29:372-6.

38. Atzeni F, Bongiovanni S, Marchesoni A, *et al*. Predictors of response to anti-TNF therapy in RA patients with moderate or high DAS28 scores. *Jt Bone Spine* 2014;81:37-40.

39. Carvalho PD, Duarte C, Vieira-Sousa E, *et al*. Predictors of response to TNF blockers in patients with polyarticular psoriatic arthritis. *Acta Reumatol Port* Published Online First: 17 August 2016.

40. Severs M, Petersen RE, Siersema PD, *et al*. Self-reported Health Care Utilization of Patients with Inflammatory Bowel Disease Correlates Perfectly with Medical Records. *Inflamm Bowel Dis* 2016;22:688-93.



7

Down the line from genome-wide association studies in inflammatory bowel disease: the resulting clinical benefits and the outlook for the future

LM Spekhorst*, MC Visschedijk*, RK Weersma and EAM Festen

*Authors contributed equally

Expert Rev Clin Immunol 2015 Jan;11(1):33-44

Abstract

Inflammatory bowel disease (IBD), consisting of Crohn's disease and ulcerative colitis, is a chronic inflammatory disease of the gut. The aetiology of IBD is complex, involving genetic as well as environmental factors. Genetic studies have identified 163 genetic risk loci for IBD, which have led to new insights into the biological mechanisms of the disease. The currently known IBD risk loci show an almost 75% overlap with genetic risk loci for other immune mediated diseases. Current studies are focused on the translation of the identified risk loci to clinical practice. The first steps towards this translation are being taken with the identification of genetic risk factors for drugs toxicity, specific disease course and response to therapy. In this review we will discuss how the IBD genetic risk loci were identified and how this knowledge can be translated towards clinical practice.

Introduction

Inflammatory bowel disease (IBD) is a chronic immune mediated disease affecting the gastrointestinal tract. The prevalence of IBD in the Western World, is approximately 1 in 1000 individuals and there is an increasing trend in incidence and prevalence in developing countries.[1] IBD consists of two distinct diseases; Crohn's disease (CD) and ulcerative colitis (UC), which have some overlapping clinical and pathological features. In CD the inflammation can occur throughout the entire gastrointestinal tract, the inflammation can affect all mucosal layers and can be complicated by strictures, and formation of abscesses and fistula. In UC the inflammation is limited to the colon and only affects the upper mucosal layer. Formation of abscesses and fistula does generally not occur in UC and stenosis is a very rare complication in UC.[2,3]

IBD is a complex disease, meaning that its aetiology is multifactorial: genetic, epigenetic and environmental factors interact and give rise to the disease. Environmental factors like smoking, medication, appendectomy, exposure to pollution, and diet have been implicated to play an essential role in the pathogenesis of IBD.[4] Smoking and prior appendectomy have been proven to be protective in UC, but in contrast smoking can aggravate inflammation in CD and increase the risk for CD.[5] The composition of the gut micriobiota, is also likely to be a major factor in IBD disease pathogenesis.[6-9]

Over the preceding decades, co-occurrence and familial aggregation of IBD was observed indicating that IBD has a strong hereditary component.[10] As the field of genetic research evolved rapidly it revealed new insights into pathogenesis of IBD. Initially genetic linkage studies were performed in families with an extremely high prevalence of IBD, which led to the identification of the first genetic risk locus for IBD, containing the gene *NOD2*.[11-13] As genome-wide-associations studies (GWAS) became available, hypothesis free testing of common genetic variants for association to disease became possible.[14] GWAS were very successful in IBD: within four years 99 genetic risk loci for IBD had been identified, making it the most successfully GWA studied immune-mediated disease.[15] Subsequently a new genotyping chip was designed which targeted areas on the genome likely to play a role in immune mediated disease: the Immunochip. Again this effort has been very successful in IBD: 163 genetic risk loci have been identified to date.[16]

The genetic risk loci identified for IBD so far have shed new light on the biological pathways underlying the disease. The translation of all of this knowledge on the genetic background of IBD towards clinical practice is still difficult. The steps from GWAS identified genetic disease risk variants to clinical practice are outlined in Figure 1. Currently projects are being undertaken to clarify the exact effect of genetic risk variants on the phenotypic variation seen in IBD. A better understanding of how genetic risk variants lead to different disease phenotypes can help us predict disease course per individual, which in turn can help us determine the most adequate treatment for each individual.

In this review we will first consider the clinical aspects of IBD. We will then highlight the new insights that GWAS and the Immunochip have brought us for IBD. We will discuss the overlap between IBD genetic risk loci and their underlying biological pathways and the genetic risk loci and pathways for other immune mediated diseases. Finally, we will discuss how these pathways can be used for therapeutic targeting and how this can help pave the way towards 'personalized medicine'.

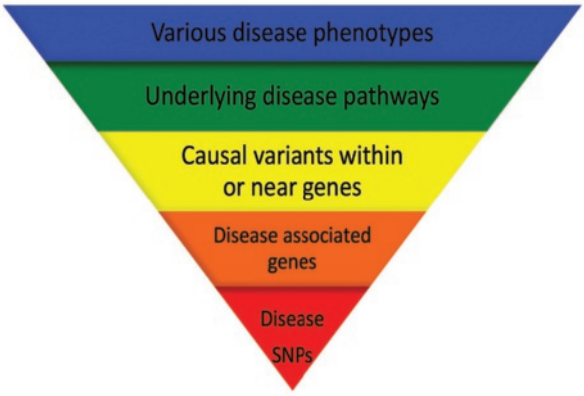


Figure 1 The steps leading from findings of genome-wide association studies (GWAS) to clinically significant outcomes. From bottom to top: In red: GWAS identify single nucleotide polymorphisms (SNPs) that are associated to the dis- ease. In orange: the next challenge is the identification of the gene that correlates with the associated SNP. In yellow: then causal variants can be identified that influence the disease associated genes. In green: once it is known which genes are involved in the disease underlying disease pathways can be constructed through co-expression and protein-protein interaction analysis. In blue: finally different genetic backgrounds can be identified that lead to slightly different disease mechanisms resulting in diverse disease phenotypes.

The clinical presentation of IBD

IBD is a chronic mucosal inflammatory disease of the gastrointestinal tract, characterized by periods of remission and relapse. Because of this dynamical character of the disease, patients can experience severe symptoms during an exacerbation, while symptoms can be mild or absent during remission. Active IBD can cause symptoms like abdominal discomfort, diarrhoea, weight loss, rectal bleeding and fatigue. In addition to inflammation of the gut, 25% of the patients have extraintestinal manifestations (EIM). Arthralgia is the most common EIM, but ophthalmological and primary mucocutaneous EIMs are also common.[2,17,18] Direct disease symptoms and EIMs influence psychosocial functioning and might result in a significantly lower quality of life and loss of work productivity.[19,20]

Disease remission can be induced and sustained by medical treatment such as mesalazine, corticosteroids, and immunosuppressants like azathioprine and anti-TNF antibody therapies.[21] Nevertheless, up to 20% of the UC patients and almost 50% of the CD patients will need surgery

within 10 years after diagnosis because of refractory disease, fibrostenotic disease, complications or development of colorectal carcinoma.[22]

IBD therapy is complicated by the fact that IBD is a heterogeneous disease with a variety of clinical phenotypes, each of which require specific treatment. Currently the treatment paradigm of IBD is shifting from treating symptomatic patients (the 'step-up' approach) to starting intensified treatment regimens early in the disease to prevent complicated disease behaviour and flares of the disease (the 'top-down' approach).[23] Before starting this top-down treatment it is essential to select those patients at risk for severe disease, to avoid high costs, 'over treating' and potentially severe side effects. At this moment one of the major issues that clinicians face is the selection of these patients at risk for a severe disease course. There are no clinical parameters or biomarkers which predict how the disease will develop, so selecting these high-risk patients is complex. Some known risk factors like early age of onset, familial occurrence of IBD and extensive disease at presentation are considered to be predictive for severe disease.[24,25] However, these factors offer a relatively slim foundation for aggressive treatment with expensive drugs with potentially severe side-effects.[26] Better understanding of the influence of environmental, genetic, and microbiomic factors on IBD phenotype will not only improve choice of treatment for IBD patients but will also enhance our understanding of the underlying disease mechanisms.

Genetic studies prior to the genome wide association scans

As mentioned previously, family studies showed that genetic factors play an important role in IBD risk: the occurrence of CD and UC in first-degree family members is respectively 10-fold and 4-fold higher compared to the general population and in monozygotic twins the concordance of CD and UC is respectively 30 and 15%.[10]

The first genetic studies in IBD were linkage and candidate gene association studies. Linkage studies map genetic risk loci by testing a series of marker alleles for co-segregation (linkage) with disease status through different generations in a family. In candidate gene association studies a candidate gene is selected based on what is known about the disease mechanism. This gene is then tagged with common genetic variants, and tested for association to a trait. A total of 10 linkage and association studies for IBD were performed in the period from 1996 to 2004.[27] Compared to similar studies in Mendelian diseases the yield of these studies in complex disease might be considered disappointing. This is mainly caused by the fact that the effect size of genetic risk variants for complex disease tends to be much smaller than that for Mendelian disease. Nonetheless, linkage studies identified the first risk gene for CD: *NOD2* on chromosome 16.[11-13] In 2001 three low frequency coding mutations (R702W, G908R and 3020insC) in the *NOD2* gene were found to be independently associated with CD in Caucasian patients.[12,13,28] These three variants lead to odds ratios (ORs) for CD between 2 and 4 in individuals heterozygote for the variants, and to ORs

between 20 and 40 in individuals homozygote for the variants. Nowadays *NOD2* variants remain the strongest genetic risk variants for CD.

Alongside *NOD2*, linkage studies suggested a link between IBD and three other genetic loci with a lower effect size: the IBD3 locus (in the *HLA*-region), the IBD5 locus (containing *SLC22A4*, *SLC22A5* and other genes), and a locus on chromosome 5q31 (containing no genes; a 'gene desert').[27]

Genome wide association studies

In the early 2000s tremendous technological advances and the progress of the Human Genome Project provided the opportunity to perform GWAS.[29,30] GWAS typically focus on associations between complex traits and 100,000–500,000 single-nucleotide polymorphisms (SNPs), selected to tag a maximum of genetic variation over the whole genome. A SNP is a DNA sequence variation, which occurs commonly in a population. A GWAS looks for statistically significant differences in allele frequencies of these SNPs between a large number of individuals with disease status (cases) and healthy controls. The associated SNPs mark genomic loci (which can contain several genes) in the human genome, which influence the risk of disease. Unlike linkage studies, GWAS are not restricted to sibling pairs and families, and thus have more statistical power to detect genetic risk loci of small to moderate effect sizes. Due to correction for multiple testing, there is a strict protocol for replication and the genome-wide statistical significance association for true positives was set to a P value < 5 10^{-8}.[31]

The first GWAS study in European ancestry CD patients confirmed *NOD2* as a CD risk gene. Moreover, it identified the association between CD and a locus containing *IL23R*, encoding a pro-inflammatory cytokine, which stimulates T-cells towards chronic inflammation.[32] The most remarkable pathway discovered through early CD GWAS studies is the autophagy pathway, which was discovered through associations at loci containing *ATG16L1* and IRGM.[33-35] The early GWAS studies in UC showed substantial overlap of genetic risk background with CD (*IL23R*, *IL12B*, *NKX2-3* en *MST1*), but also some UC specific loci (*IL10*, *HLA*). *NOD2*, *ATG16L1* and *IRGM* remain CD specific loci.[36-39]

A GWAS typically uses approximately 500–2000 cases and a similar number of controls genotyped at 100,000–500,000 SNPs. To increase the power of GWAS to detect more genetic risk variants a meta-analysis of all previously published CD GWAS was performed, which included 3000 cases and 5000 controls. This meta-analysis confirmed 11 previously putative loci and helped discover 21 new CD risk loci, including loci containing *STAT3* and *JAK2*.[40]

The appreciation of the need to further increase sample sizes led to increased collaboration through the International IBD Genetics Consortium (IIBDGC)[41] to bring together investigators and GWAS datasets from IBD genetics groups around the world. The IIBDGC published three GWAS

meta-analyses between 2008 and 2011. A CD meta-analysis of six GWAS with a total sample size of ~50,000 individuals identified 30 new loci, bringing the total count of CD genetic risk loci to 71.[42]

Similarly a meta-analysis of six GWAS studies in UC with a sample size of ~17,000 individuals identified 29 new UC loci. This increased the number of UC genetic risk loci to 47.[15]

These GWAS meta-analyses increased the total number of confirmed IBD risk loci to 99, including at least 28 shared association signals between CD and UC.

Immunochip

As GWAS results for other immune mediated diseases followed, it became clear that some risk variants were disease-specific, but that most risk variants were shared between IBD and various other immune mediated diseases. In 2011 approximately 51 of the thus far identified 99 IBD risk loci turned out to be shared with 23 different diseases, most of which immune mediated diseases.[43,44] This concept formed the basis for the development of the Immunochip: a chip composed of all genetic variants correlated to immune mediated disease.[45,46] The Immunochip covers almost 200,000 SNPs for 12 distinct immune mediated diseases (CD, UC, autoimmune thyroid disease, ankylosing spondylitis, celiac disease, IgA deficiency, multiple sclerosis (MS), primary sclerosing cholangitis (PSC), primary billiairy cirrhosis (PBC), psoriasis, rheumatoid arthritis (RA), systemic lupus erythematosus (SLE) and Type 1 diabetes (T1D)). The Immunochip was designed to densely genotype immune mediated disease risk loci with common genetic variants.[45] The Immunochip project had two main goals. The first goal was to validate the already identified disease risk loci and to establish previously putative genetic risk loci as definite genetic risk loci by testing them in a large number of new cases and controls, a process termed 'deep replication'. To achieve this goal the top ~3000 associated SNPs for each disease known from GWAS and GWAS meta-analysis were tested in a large number of case and control samples that had not been tested in previous GWAS. The second goal of the project was to fine map each risk locus to identify the most likely focus of the genetic variant that is actually causal to the disease association. To achieve this goal each known genetic risk locus was densely covered with SNPs on the Immunochip.[46]

The Immunochip, though purposefully designed, has a few limitations that should be taken into consideration.

First of all, not all loci are densely covered; putative genetic risk loci are generally only covered by a handful of SNPs. This means that, especially in the putative loci, the SNP showing the strongest association to disease is unlikely to be causal, and more likely to be in 'linkage disequilibrium' with the causal SNP. Linkage disequilibrium means that the most strongly associated SNP inherits with the causal variant because they are on a stretch of DNA that does not break during cell replication. A second limitation of the Immunochip is that it is less sensitive in non-European ethnic groups, because the SNPs have been selected from genome reference sets based on individuals of European

origin. A third limitation is that the Immunochip only includes relatively common genetic variants (minor allele frequency > 0.5%), whereas more rare variants can confer larger effects on disease risk. A final limitation, inherent to the design of the Immunochip, is that it does not cover the whole genome, but only focuses on known immune disease genetic risk loci.[46]

The number of new loci that can be identified with genome wide significance, that is with confidence that the finding is not a false positive, depends in part on the sample size and in part on the allele frequency of the genetic variants tested.[47] The Immunochip project identified 64 new risk loci for IBD increasing the number of known IBD genetic risk loci to an impressive number of 163 risk loci. Of these 163 risk loci, 30 are specific for CD and 23 for UC. The other 110 loci are shared by both diseases, implying that a shared biological mechanism plays a role in both phenotypes.[16] Whereas the Immunochip substantially increased the number of known genetic risk loci for IBD, these risk loci explain only a minority of the genetic variance in disease risk: 13–13.5% in CD and 7.5–9% in UC.[16,48] The fact that we can only explain such a small amount of risk variance in IBD suggests that other factors, like rare risk variants not tagged by the chips used so far, or interactions between genetic variants and environmental factors, also play a substantial role in the risk for IBD. Interactions within and between genetics and environment modulate the risk for a disease. These interactions are still poorly understood, which makes, for example, predicting disease risk with genetic risk variants complex.

Besides identifying new shared and unshared risk loci for several immune mediated diseases, the Immunochip also detected some highly interesting discordant associations, that is, instances in which a risk locus is shared for two diseases but where the associated variant conveys risk in one disease and is protective in the other disease. Some interesting cases of discordant associations are seen between CD and UC. Although the clinical phenotypes of the diseases are clearly overlapping, and shared risk loci are thus to be expected, a number of shared loci were found that showed a risk effect in opposite directions for each disease. An example is the locus containing the *PTPN22* gene (encoding the protein tyrosine phosphatase).[49] The main risk variant in this locus, Arg620Trp, increases risk for UC, but is protective for CD.[16,50] The mechanism underlying these discordant associations has not yet been clarified.

Immune mediated diseases

IBD is, as described earlier, a chronic inflammatory disease caused by an excessive inflammatory reaction to the host's own gut microbiome. This makes IBD one of a large range of what can be termed immune mediated disorders: chronic inflammatory diseases caused by an inflammatory reaction to an antigen that, in healthy individuals, is tolerated or quickly disposed of. For some of these diseases the antigen is known, such as the gluten protein in celiac disease. Celiac disease represents a special case of an immune mediated disorder, because the antigen can be avoided with

a gluten-free diet, thereby more or less 'curing' the disease. IBD also forms a special case among the immune mediated diseases: the instigative antigen is generally assumed to be the commensal flora of the gut. However, unlike gluten, the commensal flora of the gut is not something one can avoid. For most other immune mediated disorders, the instigative antigen is unknown, complicating the unravelling of the disease mechanism. For RA, ankylosing spondylitis, SLE, T1D and autoimmune inflammatory disease of the thyroid, antigens have been proposed but not conclusively proven.

Although each of these diseases have different phenotypes and might have different instigative antigens they share common inflammatory pathways. Immune mediated disorders are known to co-occur within families or even within individuals, suggesting they also share a common genetic background.

In 2008, early in the GWAS era, results of the different immune mediated disorders already revealed 23 genetic risk loci to be shared by two or more immune mediated diseases (ankylosing spondylitis, asthma, auto-immune thyroid disease, coeliac disease, MS, psoriasis, RA, SLE, T1D, CD and UC).[44] As data accumulated, in 2011 approximately 51 IBD genes were found to be shared with 23 different diseases, including immune mediated disorders, infectious diseases and other gastro-intestinal disorders.[43,51] Noteworthy among these shared genetic loci are the risk loci that encode proteins from the adaptive immune system (*IL23R, IL10, IL12B, IL27, IL18RAP*), these loci seem to play a role in a shared pathway for CD, UC and several other immune mediated diseases.[43]

The Immunochip once again showed that shared genetic risk loci might not have the same effect in each disease. Besides the previously observed correlated and discordant associations (the same haplotype is protective in one disease but increases the risk for another disease), it showed association signals that are non-correlated (different risk haplotypes are seen in shared risk loci), and association signals that are correlated and concordant (a risk variant increases the risk for more than one immune mediated disease).[50] This phenomenon is called pleiotropy, meaning that seemingly unrelated phenotypes can be derived from the same risk variants and that seemingly related phenotypes can derive from divergent risk variants. The mechanism underlying this pleiotropy might be that the combination of different genetic risk variants and environmental factors determine the immune mediated disease that a patient will develop.

Almost three-quarters of the IBD loci were found to be overlapping with other immune mediated diseases[43,50] and 71 loci were associated with two or more immune mediated diseases (IBD, ankylosing spondylitis, coeliac, psoriasis, RA, Type 1 Diabetes). Of these 71 loci 45% resulted in an increased risk, 14% had opposites effects (protective versus risk variant) and 42% shared the same loci but with different risk haplotypes.[50]

Pathogenetic pathways in IBD

GWAS and Immunochip meta-analysis have revealed 163 risk loci for IBD, most of which are shared for CD and UC and which identify key regulating pathways for both diseases. To advance our understanding of the mechanisms underlying disease we have to carefully select and study candidate gene(s) within each locus to see how these contribute to disease susceptibility. The most replicated and confirmed loci have been extensively studied and their candidate genes and their pathways reveal possible disease mechanisms for CD and UC also in functional studies and mouse models.

One of the strongest susceptibility loci in CD is the locus containing *NOD2*.[12,13] *NOD2* is located on chromosome 16 and encodes an intracellular pattern-recognition receptor of the innate immune system.[52,53] This receptor recognizes viral and microbial components; maintaining the gut mucosal barrier through regulation of microbiome homeostasis and activation of the innate immune response.[54] One of the mechanisms through which *NOD2* regulates microbiome homeostasis is by the production of antibacterial defensins.[55] *NOD2* receptor signalling against microbial components depends on its intracellular localization. *NOD2* risk variants cause a disrupted receptor, which make the receptor unable to recognise intracellular bacteria. This probably leads to dysbiosis of the intraluminal contents, thereby causing an inappropriate immune response.[13,56,57] Carriage of two *NOD2* risk variants also increases the risk of a severe CD disease phenotype resulting in penetrating and/or stricturing disease.[58]

Another important pathway in the disease pathogenesis of CD that has been discovered through genetic studies is autophagy (*ATG16L1*, *SMURF1*, *LRRK2* and *IRGM*).[33,59,60] Autophagy describes a cellular pathway in which organelles and foreign proteins are being delivered to the lysosome in the cell for degradation, which makes it a crucial immune defence mechanism.[61,62] Moreover there is a functional link between *ATG16L1* and *NOD2*, as *NOD2* recruits *ATG16L1* to the plasma membrane at the bacterial site of invasion to initiate autophagy. Several risk variants (SNPs), associated with *NOD2* and *ATG16L1* have been found to affect bacterial autophagy. This implies that deficient bacterial autophagy plays a key role in the disease pathogenesis of CD.[59,63,64] The autophagy related *NOD2* and *ATG16L1* variants are specific to CD, but other autophagy related genes are associated with both CD and UC (*SMURF1*, *LRRK2* and *IRGM*).[16,65]

The *IL23R* genetic IBD risk locus is part of an important disease pathway for IBD: T-helper 17 (Th17) signalling. The *IL23R* gene in this locus encodes an IL23 receptor whose subunit interacts with interleukin 23 (IL23) a pro-inflammatory cytokine. IL23 regulates the immune response against exogenous antigens in the gut, by inflammation through the production and differentiation of Th17 lymphocyte cells.[66,67] Mutations or variation in or near the *IL23R* gene are hypothesized to cause an inappropriate immune response to the commensal flora of the gut.[32,68,69] Loci with genes encoding proteins with functions downstream in the IL23-Th17 pathway have also been identified as IBD risk

loci. These genes affect components of the IL23 pathway that are expressed in Th17, Th1 and other innate lymphoid cells. Among these IBD risk loci are loci containing *JAK2* (Janus kinase 2) and *STAT3* (signal transducer and activator of transcription 3). The activation of the IL23R at the cell surface activates a secondary intracellular signalling pathway JAK2/STAT3. This *JAK/STAT* pathway plays a role in the innate and adaptive immunity, and particularly in the progression of inflammation in the Th17 cell pathway.[70] Several genes encoding pro-inflammatory cytokines in the Th17 pathway, which have been shown to be overexpressed in IBD, have also been implicated in the genetic background of IBD: *IL22*, *IL21* and *IL26*.[71-76]

'Ten-year' review: genetic research translated to the clinic

In this post-GWAS era, our knowledge on the molecular background of IBD has progressed to such a high level, that an inevitable and crucial question arises: How do we translate this knowledge to clinical practice to improve treatment of IBD or even to detect IBD early and prevent it from progressing to a full blown inflammatory disease (Figure 2)?[77]

Screening individuals at risk

Could we screen the general population for risk of IBD with the genetic risk variants that have been identified? The answer is that we could, in theory, but the predictive value of such a genetic screening test would be very low, because, as mentioned earlier, the genetic variants identified so far explain less than 20% of the genetic risk for IBD.[39,78] To improve the predictive power of this predictive test by enriching it with environmental risk factors is also not yet feasible, since our knowledge on environmental risk factors for IBD is still relatively poor. So screening the general population for IBD risk is not yet possible. Moreover, one could wonder whether we really should screen the general population for a disease that we cannot (yet) prevent. Genetic screening of individuals at high risk for IBD, for example children from families with a high prevalence of IBD with the currently known IBD risk loci, is similarly unlikely to be successful. The currently known genetic risk variants for IBD are genetic variants that are common (minor allele frequency > 1%) in the general population, and hence will not provide a test with a high specificity. Moreover, in families in which IBD is highly prevalent the underlying genetic risk variants are likely to be rare variants that may well have not yet been identified. Finally screening in high-risk individuals would only be useful if minimally invasive or preventative therapy such as dietary intervention would be available. Unfortunately, dietary interventions have not been proven effective in IBD, so early detection of IBD would only lead to earlier recognition of disease, followed by standard treatment.[79]

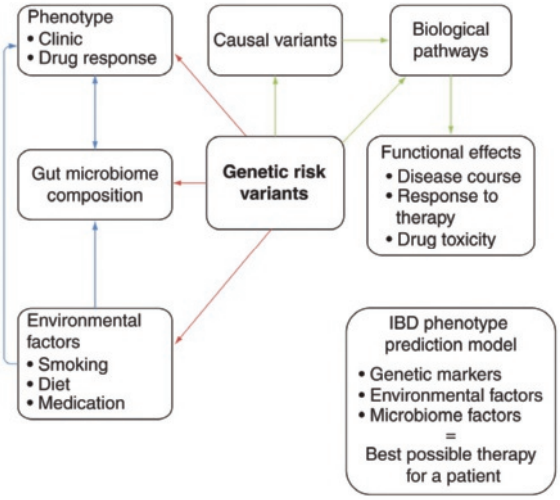


Figure 2 From genetic risk variants to the clinic. The red arrows show the interaction of genetic risk variants with other factors leading to inflammatory bowel disease (IBD) phenotypes. De blue arrows show the interaction of the non-genetic IBD risk factors with IBD phenotype. De green arrows show the route research will have to take down the line from the identification of genetic risk variants to clinical practice. The square in the lower right corner of the figure shows how our current knowledge could lead to 'personalized medicine' i.e., the best possible therapy for a patient.

Drug development

While our knowledge on the genetic background of IBD cannot be used for risk prediction, it does provide important insights into the disease mechanism of IBD. Each genetic risk locus is a potential drug target for IBD therapy. Before such targeted drugs can be developed, the target loci have to meet several important criteria.

The first criterion for the development of targeted drugs is that the causal variant in the genetic locus has to be known and the effect of this causal variant on the biological function of the gene it affects has to be understood.[80] Currently, we cannot yet meet this criterion: the genetic variants that have been identified by GWAS are variants with a high prevalence in the general population. Hence these variants are very suitable for mapping genetic variance between cases and healthy controls but they are unlikely to have a direct effect on the function of a gene. This means that the 163 genetic variants that are associated to IBD are not the true causal genetic variants that affect the genes in these loci but that these 163 variants are correlated to the causal genetic variants in these loci. Fine-mapping of these genetic risk loci and targeted re-sequencing of genetic risk loci in large cohorts of IBD cases and healthy controls is currently being performed and will lead to better understanding of the causal link between the IBD associated common genetic variants and the disease mechanism (Figure 1). The translation from common genetic risk variant to clinically significant output initially involves the identification of the gene that the variant is likely to be

correlated to by gene prioritizing. The causal gene is likely to have a differential expression as a result of the genetic variant, i.e., by identifying expression Quantitative Trait Loci (eQTL). The causal gene might also be identified because it encodes proteins that are known to interact with proteins known to be involved in IBD mechanisms, i.e., by protein-protein interaction (PPI) or through Gene Relationships Across Implicated Loci (GRAIL).

The second criterion for developing targeted drugs is that the causal genetic variant, or its effect on gene function, is drugable, and that targeting this mechanism does not lead to adverse events.[80] Some CD risk variants might for example cause an inadequate innate immune response, but upregulating this innate immune response with a drug might lead to increased inflammation caused by an exaggerated innate immune response. Other CD risk variants are located in gene deserts where the identification of a drugable candidate gene might be impossible.

In short, the road from the currently identified IBD genetic risk variants to the development of new IBD drugs seems long and hardly cost-effective. In spite of this there are several examples of approved drugs for which genetic studies identified the drug target to be associated with IBD, while the drug had already been developed independently from genetic knowledge (Table 1). Ustekinumab (Janssen-Cilag), a human antibody against IL-12 and IL-23, has for example been shown to be effective in the treatment of CD.[81,82] Ustekinumab blocks binding of IL-12 and IL-23 and thereby blocks the inflammatory cascade downstream of these interleukins. The importance of this inflammatory pathway had already been observed in the genetic background of IBD since *IL23R* and *IL12B* are important risk loci for IBD.

GWAS identified genetic loci containing *JAK2*, *STAT3* and *TYK2* genes as risk loci for IBD. The proteins encoded by these genes, *JAK2*, *STAT3* and *TYK2*, play crucial roles in secondary pro-inflammatory signalling after activation of the IL23 pathway, as described earlier. Tofacitinib (Pfizer) a selective JAK-inhibitor seems to be effective in the induction and maintenance of remission of IBD and is currently being tested in Phase III trials.[83,84]

The autophagy pathway was identified as an important disease mechanism for CD through the recognition of *ATG16L1*, *NOD2* and *IRGM* as CD risk genes. The genetic risk variants in these loci seem to lead to less effective autophagy and consequently lead to less efficient disposal of invasive microbes. Everolimus (Novartis) and sirolimus (Pfizer) are both mammalian target of rapamycin (mTOR) inhibitors and up-regulate the autophagy pathway.[85,86] Both drugs are registered for immune suppression after solid organ transplantation. Because of their known immunosuppressant effects and their specific effect on the autophagy pathway the drugs were tested in CD. Although case-reports of the treatment of CD with sirolimus seemed promising, a randomized case-control study with everolimus was terminated early because the drug showed no efficacy.[85,86] One could speculate that in the future such trials should be repeated but then including only cases with impaired autophagy e.g., carrying the *ATG16L1*, *NOD2* or *IRGM* risk variants.

Finally the homing of leukocytes to the gut seems to be an interesting disease mechanism since genetic risk loci containing *CCR6* and *CXCR5*, genes encoding proteins involved in this process, are associated to IBD.[16] Before this genetic knowledge was available drugs that impede leukocyte migration to the gut had already been developed. Initial trials of natalizumab (Biogen), an α-4 integrin inhibitor, showed promising results for inducing and maintaining remission in CD.[87,88] However serious side-effects were observed: the drug also prevents migration of leukocytes to the central nervous system, which increases the risk of severe infections of the central nervous system.[89] Vedolizumab (Millennium Pharmaceuticals, Inc) is a more specific integrin inhibitor: it specifically inhibits α-4-β-7 receptors, which makes this drug a specific inhibitor of leukocyte migration to the gut. Vedolizumab was shown to be effective for induction and short-term maintenance of remission of disease in UC patients and CD patients with prior failure on anti-TNF therapy.[90-92]

Table 1 New inflammatory bowel disease drugs based on known genetic risk factors.

Potential drugs	IBD risk genes targeted	Effect	Study status
Ustekinumab	Blocking binding of IL12-IL23	Blocks the inflammatory escades down stream	Effective in CD 82,83
Tofacitinib	Blocking selective JAK2 inhbitor	Blocking secondary pro-inflammatory signalling after activation of the IL23 pathway	Phase III trial 84,85
mTor inhibitor	ATG16L1, NOD2 and IRGM (autophagy)	Up-regulation of the autophagy pathway	Showed no efficacy 86,87
Denosumab	Apoptosis regulatory gene TNFSF11	Regulation T-cell/dendritic communication	Candidate for testing therapy 88,89

This table lists new Inflammatory Bowel Disease (IBD) drugs, the genetic risk locus (indicated with their most likely candidate gene) that they target, their presumed effect and the current status of their studies.
ATG16L1: Autophagy related 16-like 1; CD: Crohn's disease; IL12: Interleukin 12; IL23: Interleukin 23; IRGM: Immunity-related GTPase family M protein; JAK2: Janus kinase 2; NOD2: Nucleotide-binding oligomerization domain-containing protein 2; TNFSF11: Tumour necrosis factor (ligand) superfamily, member 11.

Drug repositioning

As mentioned earlier, developing new drugs targeted on IBD genetic risk loci will be a long and difficult process. The case of the testing of the mTOR inhibitors everolimus and sirolimus in CD however perfectly illustrates an alternative scenario: identifying alternative or refined indications for existing drugs approved for other indications. This process, called drug repositioning, is likely to become an important alternative to classic drug development because of the increasing costs of the development of new drugs and our increasing understanding of the biological background of diseases. For drug repositioning one considers the genetic background and biological pathways

involved in a disease and identifies registered drugs or known investigative new drugs that target these biological pathways. A new and interesting candidate for drug repositioning in IBD is denosumab. Denosumab (Amgen/GlaxoSmithKline) is registered for prevention of fractures caused by osteoporosis in postmenopausal women.[93,94] The drug acts through TNFSF11, encoded by the *TNFSF11* gene that has previously been identified as a risk locus for CD and bone mineral density.[42,95]

Predicting drug toxicity

Already important discoveries have been made in identifying genetic variants that predict drug toxicity. Variants in the Human Leukocyte Antigen (HLA) region were found to be associated with thiopurine induced pancreatitis. Patients being homozygous for the HLA-DQA1*02:01–HLA-DRB1*07:01 haplotype have a risk of 17% for developing pancreatitis after thiopurine administration. Another genetic variant that confers susceptibility to drug toxicity lies in the *NUDT15* gene, which is associated with thiopurine-induced leukopenia.[96,97] As the genetic risk variants for severe side effects often have a large effect size it is relatively easy to gain enough power, i.e., to collect a dataset large enough, to identify them. Also the clinical significance of these side effects is so big that testing for the genetic risk variant before starting a drug might be feasible. The International Serious Adverse Events Consortium (iSAEC) together with the IIBDGC are leading the research into the genetic background of the major side effects of IBD medical therapy: pancreatitis through thiopurines, kidney damage through mesalazine and neurological side-effects in anti-TNFα therapy.[98]

Five-year view & expert commentary

In recent years the treatment paradigm in IBD has changed from treating symptomatic patients to starting intensified treatment regimens early in the disease to prevent complicated disease behaviour and flares of the disease.[23] However, as mentioned before one major clinical issue in IBD is that there are no clinical parameters or biomarkers to predict how the disease will develop, so it is extremely difficult to identify patients who will benefit from aggressive treatment. Another important factor, which makes the selection of patients at risk for a severe disease course difficult, is the extreme heterogeneity of the disease. IBD has a variety of disease subphenotypes of which severity is difficult to classify (mild, moderate or severe disease). It has already been established that CD patients with a severe disease course (operations, age of onset below 40 years) carry more genetic risk variants than CD patient with a mild disease course.[78] But to identify genetic variants that are associated with specific disease phenotypes or behaviour is essential to collect clinical characteristics in a uniform and reproducible manner. So far, studies for genetic associations to subphenotypes have been performed in very small cohorts of well-phenotyped individuals or larger cohorts of poorly phenotyped individuals. Hence, only genetic risk variants with a strong risk effect have been reported to be associated to specific disease phenotypes. The *NOD2* genetic

risk variants have been reported to predict severe disease course in CD and it has been suggested that the heterogeneity of the association signal from the *HLA*-region might be caused by specific *HLA* variants being associated to specific subphenotypes of IBD.[16,42,99,100]

The Dutch government has funded a very large and exhaustively phenotyped prospective cohort of IBD patients initiated by the Parelsnoer Institute and the Initiative on Crohn and Colitis (ICC). The ICC is an initiative formed by gastroenterologists from all eight University Medical Centres of the Netherlands. The Parelsnoer Institute, financed by the Dutch government, is developing a biobank in which biomaterial en phenotypical data is being collected in a uniform matter.[101] This biobank has been funded in 2007 and already contains a large collection of both phenotype data and biomaterials. Currently, phenotype and genotype data are being integrated in an attempt to translate the genetic findings to the clinic. We hope that this integration of genetic and clinical data will expand our current knowledge on biological pathways and reveal new clinical insights.

Besides the extensively growing knowledge on the genetics of IBD, developments in gene-sequencing technologies, as well as increased availability of computational biology, have lead to novel insights into the microbial composition of the human gut microbiota. Profiling studies of the intestinal microbiome have shown that IBD is associated with characteristic shifts in the composition of the intestinal microbiota, reinforcing the view that IBD results from altered interactions between intestinal microbes and the mucosal immune system.[102,103] Decreased complexity of the gut microbial ecosystem is a common feature in patients with CD or UC.[9,104] The human microbiome is a very dynamical and interactive system. Future studies with a multifaceted approach to the microbiome in IBD are essential. From a clinical perspective the increased understanding of the microbiome will hopefully lead to new treatment options like anti- and probiotics.[9,93]

Once we have established which factors predict severe disease outcomes in IBD we might be able to start intensive treatment early on in disease in patients with predicted severe disease and spare patients with predicted mild disease excess treatment and unnecessary side-effects.[23]

Conclusion

In this review we have outlined the development of genetic studies in IBD from linkage studies, via GWAS, to the Immunochip study. We have discussed the shared genetic and biological background of IBD with other immune mediated diseases. We have outlined what genetic studies have taught us about the disease mechanisms underlying IBD. Finally we have discussed how these findings can be translated to clinical practice.

GWAS and the Immunochip have shown us that the underlying predisposition for the IBD phenotype is diminished innate immune response followed by an exaggerated inflammatory reaction to the commensal flora of the gut. However, genetic variants by themselves explain only a small portion of disease risk. We have to explore their interaction with environmental factors and

their association to specific disease phenotypes in order to gain a comprehensive understanding of disease mechanisms. Currently large prospective and retrospective studies are being performed focused on identifying genetic risk variants that can predict a specific disease course, response to therapy, or development of drug toxicity. Only through such large well-phenotyped multi-omics studies, will we be able to translate our knowledge on the genetic background of IBD to clinical practice. At the outset the main benefit of translating genetic risk variants to clinical practice will be the prediction of drug toxicity or severe side effects from IBD therapy. The first genetic variants predicting drug toxicity are currently being published. In the near future we hope that our knowledge on the genetic background of IBD can be the basis for the development of targeted drug therapies. The process of drug repositioning based on genetic knowledge on IBD could lead to quick wins in the development of targeted therapy, and this venue should be pursued. A final important promise that our knowledge on the genetic background of IBD holds is that of personalized medicine: adapting treatment to each individual patient based on his or her genetic profile. Large studies with multi-omics data on each patient are being performed to realise the data integration needed in order to achieve personalized medicine.

We hope that in a few years we can predict disease course and best possible treatment for each patient at diagnosis with a predictive test constructed of genetic markers, microbial markers, protein markers and environmental factors. By this time the range of possible IBD treatments should have increased, and we should have a wide choice in targeted treatments for severe disease, but also for the treatment of very mild disease.

References

• of interest

•• of considerable interest

1. Molodecky N, Soon I, Rabi D, *et al*. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology* 2012;142(1):46-54.
2. Bernstein CN, Fried M, Krabshuis JH, *et al*. World Gastroenterology Organization Practice Guidelines for the diagnosis and management of IBD in 2010. *Inflamm Bowel Dis* 2010;16(1):112-24.
3. Sands BE. From symptom to diagnosis: clinical distinctions among various forms of intestinal inflammation. *Gastroenterology* 2004;126(6):1518-32.
4. Molodecky NA, Soon IS, Rabi DM, *et al*. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. *Gastroenterology* 2012;142(1):46-54.e42.quiz e30.
5. O'Toole A, Korzenik J. Environmental triggers for IBD. *Curr Gastroenterol Rep* 2014;16(7):396.
6. Nagalingam NA, Lynch S V. Role of the microbiota in inflammatory bowel diseases. *Inflamm Bowel Dis* 2012;18(5):968-84.
7. Manichanh C, Borruel N, Casellas F, *et al*. The gut microbiota in IBD. *Nat Rev Gastroenterol Hepatol* 2012;9(10):599-608.
8. Leone V, Chang EB, Devkota S. Diet, microbes, and host genetics: the perfect storm in inflammatory bowel diseases. *J Gastroenterol* 2013;48(3):315-21.
9. Kostic AD, Xavier RJ, Gevers D. The microbiome in inflammatory bowel disease: current status and the future ahead. *Gastroenterology* 2014;146(6):1489-99.
10. Brant SR. Update on the heritability of inflammatory bowel disease: the importance of twin studies. *Inflamm Bowel Dis* 2011;17(1):1-5.
11. Hugot JP, Laurent-Puig P, Gower-Rousseau C, *et al*. Mapping of a susceptibility locus for Crohn's disease on chromosome 16. *Nature* 1996;379(6568):821-3.
12. Hugot JP, Chamaillard M, Zouali H, *et al*. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* 2001;411(6837):599-603.
13. Ogura Y, Bonen DK, Inohara N, *et al*. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* 2001;411(6837):603-6.
14. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447(7145):661-78.
15. Anderson CA, Boucher G, Lees CW, *et al*. Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat Genet* 2011;43(3):246-52.
16. Jostins L, Ripke S, Weersma RK, *et al*. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;491(7422):119-24.
17. Lakatos PL, Lakatos L, Kiss LS, *et al*. Treatment of extraintestinal manifestations in inflammatory bowel disease. *Digestion* 2012;86(Suppl 1):28-35.
18. Ott C, Schölmerich J. Extraintestinal manifestations and complications in IBD. *Nat Rev Gastroenterol Hepatol* 2013;10(10):585-95.
19. Turnbull GK, Vallis TM. Quality of life in inflammatory bowel disease: the interaction of disease activity with psychosocial function. *Am J Gastroenterol* 1995;90(9):1450-4.
20. Blumberg RS, Dittel B, Hafler D, *et al*. Unraveling the autoimmune translational research process layer by layer. *Nat Med* 2012;18(1):35-41.
    - Presents general model for translational research from genetics to clinical practice for immune mediated disease.
21. Mowat C, Cole A, Windsor A, *et al*. Guidelines for the management of inflammatory bowel disease in adults. *Gut* 2011;60(5):571-607.

22. Frolkis AD, Dykeman J, Negrón ME, *et al*. Risk of surgery for inflammatory bowel diseases has decreased over time: a systematic review and meta-analysis of population-based studies. *Gastroenterology* 2013;145(5):996-1006.

23. D'Haens G, Baert F, van Assche G, *et al*. Early combined immunosuppression or conventional management in patients with newly diagnosed Crohn's disease: an open randomised trial. *Lancet* 2008;371(9613):660-7.

24. Imielinski M, Baldassano RN, Griffiths A, *et al*. Common variants at five new loci associated with early-onset inflammatory bowel disease. *Nat Genet* 2009;41(12):1335-40.

25. Kugathasan S, Baldassano RN, Bradfield JP, *et al*. Loci on 20q13 and 21q22 are associated with pediatric-onset inflammatory bowel disease. *Nat Genet* 2008;40(10):1211-15.

26. Golovics PA, Mandel MD, Lovasz BD, *et al*. Inflammatory bowel disease course in Crohn's disease: is the natural history changing? *World J Gastroenterol* 2014;20(12):3198-207.

27. Mathew CG, Lewis CM. Genetics of inflammatory bowel disease: progress and prospects. *Hum Mol Genet* 2004;13 Spec No:R161-8.

28. Cuthbert AP, Fisher SA, Mirza MM, *et al*. The contribution of NOD2 gene mutations to the risk and site of disease in inflammatory bowel disease. *Gastroenterology* 2002;122(4):867-74.

29. Sachidanandam R, Weissman D, Schmidt SC, *et al*. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* 2001;409(6822):928-33.

30. Lander ES, Linton LM, Birren B, *et al*. Initial sequencing and analysis of the human genome. *Nature* 2001;409(6822):860-921.

31. Rothman KJ, Greenland S, Lash TL. Modern Epidemiology. Khoury MJ, Millikan R, Gwinn M, editors, 3rd Edition. Genetic and Molecular Epidemiology. Lippincott, Williams & Wilkins; Philadelphia, PA: 2008.

32. Duerr RH, Taylor KD, Brant SR, *et al*. A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. *Science* 2006;314(5804):1461-3.

33. Hampe J, Franke A, Rosenstiel P, *et al*. A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. *Nat Genet* 2007;39(2):207-11.

34. Parkes M, Barrett JC, Prescott NJ, *et al*. Sequence variants in the autophagy gene IRGM and multiple other replicating loci contribute to Crohn's disease susceptibility. *Nat Genet* 2007;39(7):830-2.

35. WTCCC. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* 2007;447(7145):661-78.

36. Fisher SA, Tremelling M, Anderson CA, *et al*. Genetic determinants of ulcerative colitis include the ECM1 locus and five loci implicated in Crohn's disease. *Nat Genet* 2008;40(6):710-12.

37. Silverberg MS, Cho JH, Rioux JD, *et al*. Ulcerative colitis-risk loci on chromosomes 1p36 and 12q15 found by genome-wide association study. *Nat Genet* 2009;41(2):216-20.

38. Franke A, Balschun T, Karlsen TH, *et al*. Sequence variants in IL10, ARPC2 and multiple other loci contribute to ulcerative colitis susceptibility. *Nat Genet* 2008;40(11):1319-23.

39. Festen EAM, Stokkers PCF, van Diemen CC, *et al*. Genetic analysis in a Dutch study sample identifies more ulcerative colitis susceptibility loci and shows their additive role in disease risk. *Am J Gastroenterol* 2010;105(2):395-402.

40. Barrett JC, Hansoul S, Nicolae DL, *et al*. Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease. *Nat Genet* 2008;40(8):955-62.

41. IIBDGC. Available from: www.ibdgenetics.org/

42. Franke A, McGovern DPB, Barrett JC, *et al*. Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci. *Nat Genet* 2010;42(12):1118-25.

43. Lees CW, Barrett JC, Parkes M, *et al*. New IBD genetics: common pathways with other diseases. *Gut* 2011;60(12):1739-53.

44. Zhernakova A, van Diemen CC, Wijmenga C. Detecting shared pathogenesis from the shared genetics of immune-related diseases. *Nat Rev Genet* 2009;10(1):43-55.

45. Trynka G, Hunt KA, Bockett NA, *et al*. Dense genotyping identifies and localizes multiple common and rare variant association signals in celiac disease. *Nat Genet* 2011;43(12):1193-201.

46. Cortes A, Brown MA. Promise and pitfalls of the Immunochip. *Arthritis Res Ther* 2011;13(1):101.

47. Visscher PM, Brown MA, McCarthy MI, *et al*. Five years of GWAS discovery. *Am J Hum Genet* 2012;90(1):7-24.

48. Chen G-B, Lee SH, Brion M-JA, *et al*. Estimation and partitioning of (co)heritability of inflammatory bowel disease from GWAS and immunochip data. *Hum Mol Genet* 2014;23(17):4710-20.

49. Diaz-Gallo L-M, Espino-Paisán L, Fransen K, *et al*. Differential association of two PTPN22 coding variants with Crohn's disease and ulcerative colitis. *Inflamm Bowel Dis* 2011;17(11):2287-94.

50. Parkes M, Cortes A, van Heel DA, *et al*. Genetic insights into common pathways and complex relationships among immune-mediated diseases. *Nat Rev Genet* 2013;14(9):661-73.

    • Indepth discussion of shared disease pathways between inflammatory diseases.

51. Festen EAM, Goyette P, Green T, *et al*. A meta-analysis of genome-wide association scans identifies IL18RAP, PTPN2, TAGAP, and PUS10 as shared risk loci for Crohn's disease and celiac disease. *PLoS Genet* 2011;7(1):e1001283.

52. Van Limbergen J, Radford-Smith G, Satsangi J. Advances in IBD genetics. *Nat Rev Gastroenterol Hepatol* 2014;11(6):372-85.

53. Girardin SE, Boneca IG, Viala J, et al. Nod2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection. *J Biol Chem* 2003;278(11):8869-72.

54. Couturier-Maillard A, Secher T, Rehman A, *et al*. NOD2-mediated dysbiosis predisposes mice to transmissible colitis and colorectal cancer. *J Clin Invest* 2013;123(2):700-11.

55. Kobayashi KS, Chamaillard M, Ogura Y, *et al*. Nod2-dependent regulation of innate and adaptive immunity in the intestinal tract. *Science* 2005;307(5710):731-4.

56. Philpott DJ, Sorbara MT, Robertson SJ, *et al*. NOD proteins: regulators of inflammation in health and disease. *Nat Rev Immunol* 2014;14(1):9-23.

57. Lipinski S, Grabe N, Jacobs G, *et al*. RNAi screening identifies mediators of NOD2 signaling: implications for spatial specificity of MDP recognition. *Proc Natl Acad Sci USA* 2012;109(52):21426-31.

58. Adler J, Rangwalla SC, Dwamena BA, *et al*. The prognostic power of the NOD2 genotype for complicated Crohn's disease: a meta-analysis. *Am J Gastroenterol* 2011;106(4):699-712.

59. Van Limbergen J, Kabakchiev B, Stempak JM, *et al*. Hypothesis-free analysis of ATG16L1 demonstrates gene-wide extent of association with Crohn's disease susceptibility. *Gut* 2013;62(2):331-3.

60. Rioux JD, Xavier RJ, Taylor KD, *et al*. Genome-wide association study identifies new susceptibility loci for Crohn disease and implicates autophagy in disease pathogenesis. *Nat Genet* 2007;39(5):596-604.

61. Hussey S, Travassos LH, Jones NL. Autophagy as an emerging dimension to adaptive and innate immunity. *Semin Immunol* 2009;21(4):233-41.

62. Mizushima N, Levine B, Cuervo AM, *et al*. Autophagy fights disease through cellular self-digestion. *Nature* 2008;451(7182):1069-75.

63. Travassos LH, Carneiro LAM, Ramjeet M, *et al*. Nod1 and Nod2 direct autophagy by recruiting ATG16L1 to the plasma membrane at the site of bacterial entry. *Nat Immunol* 2010;11(1):55-62.

64. Cooney R, Baker J, Brain O, *et al*. NOD2 stimulation induces autophagy in dendritic cells influencing bacterial handling and antigen presentation. *Nat Med* 2010;16(1):90-7.

65. Orvedahl A, Sumpter R, Xiao G, *et al*. Image-based genome-wide siRNA screen identifies selective autophagy factors. *Nature* 2011;480(7375):113-17.

66. McGovern D, Powrie F. The IL23 axis plays a key role in the pathogenesis of IBD. *Gut* 2007;56(10):1333-6.

67. Hue S, Ahern P, Buonocore S, *et al*. Interleukin-23 drives innate and T cell-mediated intestinal inflammation. *J Exp Med* 2006;203(11):2473-83.

68. Brand S. Crohn's disease: th1, Th17 or both? The change of a paradigm: new immunological and genetic insights implicate Th17 cells in the pathogenesis of Crohn's disease. *Gut* 2009;58(8):1152-67.

69. Yu RY, Gallagher G. A naturally occurring, soluble antagonist of human IL-23 inhibits the development and in vitro function of human Th17 cells. *J Immunol* 2010;185(12):7302-8.

70. Coskun M, Salem M, Pedersen J, *et al*. Involvement of JAK/STAT signaling in the pathogenesis of inflammatory bowel disease. *Pharmacol Res* 2013;76:1-8.

71. Brand S, Beigel F, Olszak T, *et al*. IL-22 is increased in active Crohn's disease and promotes proinflammatory gene expression and intestinal epithelial cell migration. *Am J Physiol Gastrointest Liver Physiol* 2006;290(4):G827-38.

72. Fujino S, Andoh A, Bamba S, *et al*. Increased expression of interleukin 17 in inflammatory bowel disease. *Gut* 2003;52(1):65-70.

73. Annunziato F, Cosmi L, Santarlasci V, *et al*. Phenotypic and functional features of human Th17 cells. *J Exp Med* 2007;204(8):1849-61.

74. Seiderer J, Elben I, Diegelmann J, *et al*. Role of the novel Th17 cytokine IL-17F in inflammatory bowel disease (IBD): upregulated colonic IL-17F expression in active Crohn's disease and analysis of the IL17F p.His161Arg polymorphism in IBD. *Inflamm Bowel Dis* 2008;14(4):437-45.

75. Dambacher J, Beigel F, Zitzmann K, *et al*. The role of the novel Th17 cytokine IL-26 in intestinal inflammation. *Gut* 2009;58(9):1207-17.

76. Christophi GP, Rong R, Holtzapple PG, *et al*. Immune markers and differential signaling networks in ulcerative colitis and Crohn's disease. *Inflamm Bowel Dis* 2012;18(12):2342-56.

77. Fransen K, Mitrovic M, van Diemen CC, *et al*. The quest for genetic risk factors for Crohn's disease in the post-GWAS era. *Genome Med* 2011;3(2):13.

78. Weersma RK, Stokkers PCF, van Bodegraven AA, *et al*. Molecular prediction of disease risk and severity in a large Dutch Crohn's disease cohort. *Gut* 2009;58(3):388-95.

79. Hou JK, Lee D, Lewis J. Diet and Inflammatory Bowel Disease: review of Patient-Targeted Recommendations. *Clin Gastroenterol Hepatol* 2014;12(10):1592-600.

80. Plenge RM, Scolnick EM, Altshuler D. Validating therapeutic targets through human genetics. *Nat Rev Drug Discov* 2013;12(8):581-94.

    • Presents an overview of the steps that have to be taken to go from genetic findings to drug targets.

81. Sandborn WJ, Feagan BG, Fedorak RN, *et al*. A randomized trial of Ustekinumab, a human interleukin-12/23 monoclonal antibody, in patients with moderate-to-severe Crohn's disease. *Gastroenterology* 2008;135(4):1130-41.

82. Sandborn WJ, Gasink C, Gao L-L, *et al*. Ustekinumab induction and maintenance therapy in refractory Crohn's disease. *N Engl J Med* 2012;367(16):1519-28.

83. Sandborn WJ, Ghosh S, Panes J, *et al*. Tofacitinib, an oral Janus kinase inhibitor, in active ulcerative colitis. *N Engl J Med* 2012;367(7):616-24.

84. Sandborn WJ, Ghosh S, Panes J, *et al*. A Phase 2 Study of Tofacitinib, an Oral Janus Kinase Inhibitor, in Patients with Crohn's Disease. *Clin Gastroenterol Hepatol* 2014;12(9):1485-93.e2.

85. Reinisch W, Panés J, Lémann M, *et al*. A multicenter, randomized, double-blind trial of everolimus versus azathioprine and placebo to maintain steroid-induced remission in patients with moderate-to-severe active Crohn's disease. *Am J Gastroenterol* 2008;103(9):2284-92.

86. Massey DCO, Bredin F, Parkes M. Use of sirolimus (rapamycin) to treat refractory Crohn's disease. *Gut* 2008;57(9):1294-6.

87. Sakuraba A, Annunziata ML, Cohen RD, *et al*. Mucosal healing is associated with improved long-term outcome of maintenance therapy with natalizumab in Crohn's disease. *Inflamm Bowel Dis* 2013;19(12):2577-83.

88. Juillerat P, Wasan SK, Fowler SA, *et al*. Efficacy and safety of natalizumab in Crohn's disease patients treated at 6 Boston academic hospitals. *Inflamm Bowel Dis* 2013;19(11):2457-63.

89. Yousry TA, Major EO, Ryschkewitsch C, *et al*. Evaluation of patients treated with natalizumab for progressive multifocal leukoencephalopathy. *N Engl J Med* 2006;354(9):924-33.

90. Sandborn WJ, Feagan BG, Rutgeerts P, *et al*. Vedolizumab as induction and maintenance therapy for Crohn's disease. *N Engl J Med* 2013;369(8):711-21.

91. Feagan BG, Rutgeerts P, Sands BE, *et al*. Vedolizumab as induction and maintenance therapy for ulcerative colitis. *N Engl J Med* 2013;369(8):699-710.

92. Thomas S, Baumgart DC. Targeting leukocyte migration and adhesion in Crohn's disease and ulcerative colitis. *Inflammopharmacology* 2012;20(1):1-18.

93. Lacey DL, Boyle WJ, Simonet WS, *et al*. Bench to bedside: elucidation of the OPG-RANK-RANKL pathway and the development of denosumab. *Nat Rev Drug Discov* 2012;11(5):401-19.

94. Häusler KD, Horwood NJ, Chuman Y, *et al*. Secreted frizzled-related protein-1 inhibits RANKL-dependent osteoclast formation. *J Bone Miner Res* 2004;19(11):1873-81.

95. Tilg H, Moschen AR, Kaser A, *et al*. Gut, inflammation and osteoporosis: basic and clinical concepts. *Gut* 2008;57(5):684-94.

96. Yang S-K, Hong M, Baek J, *et al*. A common missense variant in NUDT15 confers susceptibility to thiopurine-induced leukopenia. *Nat Genet* 2014;46(9):1017-20.

97. Heap GA, Weedon MN, Bewshea CM, *et al*. HLA-DQA1-HLA-DRB1 variants confer susceptibility to pancreatitis induced by thiopurine immunosuppressants. *Nat Genet* 2014;46(10):1131-4.

    • The identification of a genetic variant that predicts a severe side-effect of thiopurines. Predicting severe side-effects will be one of the first clinical applications of genetic tests in complex disease in the clinic.

98. International SAE Consortium. Available from: www.saeconsortium.org

99. Cleynen I, González JR, Figueroa C, *et al*. Genetic factors conferring an increased susceptibility to develop Crohn's disease also influence disease phenotype: results from the IBDchip European Project. *Gut* 2013;62(11):1556-65.

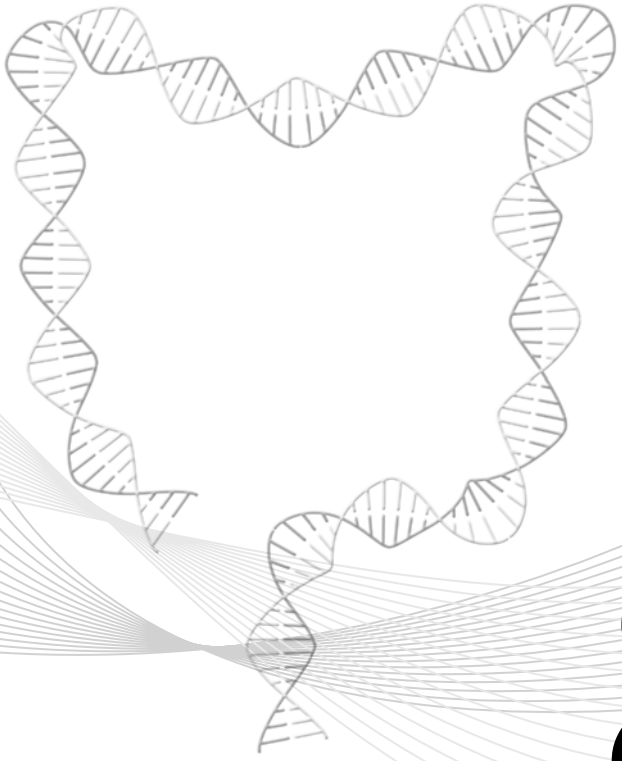
100. Cleynen I, Mahachie John JM, Henckaerts L, *et al*. Molecular reclassification of Crohn's disease by cluster analysis of genetic variants. *PLoS One* 2010;5(9):e12952.

101. PSI. Available from: www.string-of-pearls.org

102. Morgan XC, Tickle TL, Sokol H, *et al*. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. *Genome Biol* 2012;13(9):R79.

103. Gevers D, Kugathasan S, Denson LA, *et al*. The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe* 2014;15(3):382-92.

104. Huttenhower C, Kostic AD, Xavier RJ. Inflammatory Bowel Disease as a Model for Translating the Microbiome. *Immunity* 2014;40(6):843-54.



8

Identification of clinical and genetic parameters associated with hidradenitis suppurativa in inflammatory bowel disease

IC Janse,* MJ Koldijk,* LM Spekhorst, A Vich Vila, RK Weersma, G Dijkstra, and B Horváth

*Authors contributed equally

Inflamm Bowel Dis 2016 Jan;22(1):106-13

Abstract

Background: Hidradenitis suppurativa (HS) has recently been associated with inflammatory bowel disease (IBD). The objective of this study is to investigate the prevalence of HS in IBD and to identify clinical and genetic parameters associated with HS in IBD.

Methods: A questionnaire, validated for HS, was sent to 1969 patients suffering from IBD.

Results: The prevalence of HS in our IBD cohort (1260 participating patients) was significantly higher than in the general population (6.8%-10.6% versus 1%-2%). IBD patients with HS were affected by IBD significantly earlier and more often treated with anti-TNFα therapy and surgical resection compared to IBD without HS. Female gender, smoking, a higher body mass index, and younger age were independent associated parameters for HS. Within cases allelic association analysis was performed for 59 cases (IBD with HS) and 293 controls (IBD without HS). We observed 2 promising new associations in genomic regions harbouring *ELOVL7* (rsnumber 10057395 $P = 7.15 \times 10^{-5}$, odds ratio = 0.4), and in the intergenic region between *SULT1B1* and *SULT1E1* (rsnumber 2014777 $P = 7.48 \times 10^{-5}$, odds ratio = 2.3).

Conclusions: HS is present in 6.8% to 10.6% of IBD patients. Co-morbid HS is associated with an early onset of IBD in which anti-tumour necrosis factor-α therapy and surgical resections are often needed. We identified a suggestive protective association with *ELOVL7* and suggestive risk association with the genes *SULT1B1* and *SULT1E1* for HS, in the context of IBD. These genetic associations need further exploration and replication in additional independent cohorts.

Introduction

Hidradenitis suppurativa (HS) is a skin condition in apocrine gland bearing regions of the body.[1] The prevalence in Europe is thought to be approximately 1%.[2] Inflammatory bowel disease (IBD) comprises Crohn's disease (CD) and ulcerative colitis (UC).[3] In Western Europe 0.5% to 1% of the population is diagnosed with IBD.[4] There are numerous similarities between HS and IBD, raising the hypothesis for similar pathogenesis. Clinically, HS and CD, are characterized by the formation of sinus tracts and scarring.[1,5] In HS, it has been established that dilatation of the terminal hair follicle leads to rupture of the follicle with leakage of its content into surrounding tissue.[6] The significant associated inflammation is characterized by dysregulation of the immune system.[7,8] In IBD, an inappropriate immune response to commensal bacteria leads to a continuous inflammatory response.[9] Th-17 cells and tumour necrosis factor (TNF)-α are considered to be involved in both HS and IBD.[6,10] Moreover it is well known that both chronic inflammatory diseases respond to anti-TNFα therapy.[5]

CD and HS share similar risk factors. Smoking, being overweight, and female gender are the most important risk factors in HS.[11] For CD, smoking cigarettes is considered to be one of the most important risk factors as well.[12-14] Remarkably, UC patients have an increased risk of a flare if they quit smoking. Furthermore, high body mass index (BMI) appears to be a risk factor for IBD, mainly for CD.[15]

The role of genetic susceptibility in HS is a subject of ongoing research. Up to 40% of HS patients show a positive family history for the disease. Familial occurrences of HS which follow an autosomal dominant pattern of inheritance with 100% penetrance have been reported.[11] A candidate locus for HS was identified at chromosome 1p21.1-1q25.3.16. However, Ali-Ali *et al* [17] could not confirm that HS is linked to loci on chromosome 1p21.1-1q25.3. Mutations in genes encoding the component of gamma-secretase (presenilin-1, presenilin enhancer-2, and nicastrin) have also been reported.[18] Gamma-secretase is a transmembranous enzyme complex which cleaves the intracellular domain of Notch and thereby enhances intracellular Notch signalling. Notch deficient mice show occlusion of hair follicles, which is the primary event in HS.[19,20] Finally, TNF gene polymorphisms may play a role in susceptibility to HS.[21]

The genetic background is important in the pathogenesis of IBD as well. The nucleotide-binding oligomerization domain containing 2 (*NOD2*) gene was the first identified risk gene for CD.[22] In recent years, large-scale genome-wide association studies have identified 163 independent genomic loci to be associated with IBD.[23] The majority of these 163 susceptibility loci are associated with both CD and UC, suggesting that both diseases have largely overlapping biological mechanisms.

In 2010, van der Zee *et al* showed a possible association between HS and IBD in a pilot study. The prevalence of HS was 16% in their IBD population.[5] Recently, van der Zee *et al* [24] confirmed a

co-incidence in a cohort of 1093 IBD patients. Using anonymous questionnaires, they found a HS prevalence of 23% in this large group of IBD patients. The genetic basis of this association remains unknown.

The objective of this study is to investigate the association between HS and IBD in a large cohort of IBD patients. Moreover, we search for clinical parameters associated with HS in IBD and perform a genetic allelic association analysis to identify genetic variants underlying HS development in IBD.

Materials and methods

Inclusion of Patients

All patients diagnosed with IBD in the University Medical Center Groningen before February 2014 were asked to participate in this study. The diagnosis of IBD was made by a gastroenterologist based on clinical, endoscopic, and histopathological features. Patients with IBD-unclassified and microscopic colitis were excluded. The participants were requested to fill in a questionnaire (see Appendices, **Supplemental Digital Content 1 and 2**, http://links.lww.com/IBD/B125 and http://links.lww.com/IBD/B126), which was sent to their private address. After 4 weeks, a reminder was sent to the non-respondents. Anonymously returned questionnaires were excluded. The database was closed after 10 weeks. According to Dutch law and the ethical committee, a separate informed consent was not needed for sending such a questionnaire. Genotyping was performed in patients who had given written informed consent according to the Dutch Parelsnoer Institute biobanking protocol.

Questionnaire and Verification of Diagnosis of HS

The questionnaire contained questions about patient characteristics (date of birth, and gender), risk factors (smoking behaviour, family history, length, and weight), and HS. The HS questions were based on a previously validated questionnaire.[25] In addition, prototypical colour pictures of HS lesions in different stages of the disease were added, enabling the patients to self-assess their presence of HS.

Two sources of ascertaining HS were used. First, the medical records of the patients were checked. We considered the diagnosis of HS valid, if it was confirmed by a dermatologist, gastroenterologist, or surgeon, and if the exact location of the lesions were known. Secondly, in the remaining patients, verification by phone took place. Detailed information was obtained about the presence of inflammation and the location of the skin disorder.

IBD Evaluation

The data of all valid respondents was checked on the type of IBD. Moreover, the severity of IBD was assessed by determining the Montreal classification.[26] Additionally, it was established if patients were ever treated with immunosuppressive agents (azathioprine, 6-mercaptopurine, methotrexate, 6-thioguanine, cyclosporine, and tacrolimus) and/or anti-TNFα therapy and whether they underwent surgical resection in the past.

Statistical Analysis

Statistics were performed using IBM SPSS 20.0 software for Windows. Descriptive analyses were applied for all relevant variables. After performing this overall analysis, it was also executed separately for the 2 major diagnostic categories CD and UC. Group comparison was done by applying independent samples t tests or Mann-Whitney U tests for continuous data, and chi-square tests or Fisher's exact tests (if the number of subjects in a cell was less than 5) for dichotomous variables. Association between HS and IBD was examined by estimating the proportion (π) of HS within the study population. Subsequently, the prevalence of HS in the general, healthy population was compared with the prevalence of our IBD study population.

To create an association model, multivariate logistic regression analysis was performed using the presence of HS as the dependent variable. The explanatory variables were gender (male versus female), IBD type (CD versus UC), smoking (current, former, or non-smokers), BMI, and age. The odds ratios (OR) were calculated and provided by a confidence interval (CI). Finally, the likelihood ratio test was performed. Significance levels were set at $\alpha = 0.05$.

Genetic Analysis

We performed genotyping in 355 cases (59 IBD patients with HS and 296 IBD patients without HS) using the Illumina Immunochip (Illumina, Inc., San Diego, CA), which is a custom-made genotyping array designed to densely cover the immune related risk loci with common genetic variants. The Immunochip comprises of ~200,000 SNPs derived from the analysis of genome-wide association studies for 12 immune-mediated diseases.[23]

Raw intensities were normalized using Illumina's Genome Studio program (Illumina, Inc.). Clustering of the intensities and genotype calling was performed using the optiCall clustering program, with a no-calling cutoff set to 0.7.[27] To avoid false positive signals in the genome-wide association studies analysis, a stringent quality control (QC) was performed using PLINK software version 1.07.[28]

SNP QC

Single nucleotide polymorphisms (SNPs) meeting the following criteria were included in the allelic association analysis: being located in the autosomal chromosomes, a Hardy-Weinberg equilibrium test with a P value > 0.001, a call rate equal or bigger than 98% and a minor allele frequency greater than 0.05 (minor allele frequency > 0.05).

Sample QC

Samples with low genotyping efficiency were removed (call rate < 90%). To identify duplicates or relatedness in the sample dataset, SNPs in high linkage disequilibrium were removed; remaining SNPs were pruned 3 times for linkage disequilibrium ($r^2 < 0.2$), with a window size of 50 SNPs and a step size of 5. A subset of 14,618 SNPs was used to calculate the identity-by-descent in PLINK (--genome). Duplicate samples were identified by using an identity value higher than 0.8 (pi-hat > 0.8), related samples were identified by using an identity value higher than 0.4 (pi-hat > 0.4). After QC, the dataset contained 112,974 SNPs, 59 IBD patients with HS, and 293 IBD patients without HS.

Genetic Allelic Association Analysis

Within cases allelic association analysis was performed using chi-square test (--assoc) in PLINK. The results of the association were presented in a Manhattan plot using R statistical package (Package "qqman," Version 0.1.2). In addition, the genotype calling quality for SNPs with the highest P value were checked manually.

Results

Of the 1969 sent questionnaires, 652 (33%) were not returned. Five patients were excluded due to anonymous submission, 31 questionnaires were returned to sender because the addresses were invalid and 21 patients were excluded for suffering from IBD unclassified (n = 13) or microscopic colitis (n = 9) instead of IBD. A total of 1260 (64%) IBD patients replied the questionnaire well-directed and were enrolled.

The baseline characteristics of our patients are demonstrated in Table 1. There is an equal distribution of CD (634) and UC (626) patients.

Prevalence of HS in IBD

Verification of HS was executed in the 246 participants (19.5%) with a positive answer to the HS question. The diagnosis of HS was known from medical dermatology records in 24 cases. The other 222 patients were called; in 110 patients the diagnosis HS was confirmed. In total, 134 of the 1260

to 1969 IBD patients were suffering from HS. The prevalence of validated HS is thus between 6.8 (SD 0.0057% and 95% CI: 0.06-0.08) and 10.6% (SD 0.0086% and 95% CI: 0.09-0.12). With regard to CD and UC separately, HS was present in 96 of the 634 CD patients and 38 of the 626 UC patients. The 95% CI was 0.124 to 0.179 (SD 0.0142) for HS in CD and 0.042 to 0.079 (SD 0.0095) for HS in UC.

Table 1 Patient characteristics of the IBD population

	IBD total n (%)
Number of patients	1260
IBD type	
CD	634 (50.3)
UC	626 (49.7)
Gender	
Female	740 (58.7)
Male	520 (41.3)
Smoking status[a]	
Non-smokers	515 (41.1)
Former smokers	500 (39.9)
Current smokers	240 (19.0)
Pack years[a] median (range)	
Current smokers	9.0 (0.03-47.50)
Former smokers	7.5 (0.05-80.00)
Age[b] (mean years) (SD)	47.3 (15.8)
BMI (mean kg/m^2)[a] (SD)	25.1 (4.5)

[a]Missing values for smoking status (n = 5), pack years (n = 145) and BMI (n = 23).
[b]Age was determined at date of closing database (May 21st, 2014).
IBD: inflammatory bowel disease; CD: Crohn's disease; UC: ulcerative colitis.

IBD with HS Versus IBD Without HS

IBD patients with HS were significantly (P < 0.001), more likely to suffer from CD (71.6%) than from UC (28.4%) as shown in Table 2. The prevalence of women (84.3%) was significantly (P < 0.001) higher in IBD with HS compared with IBD without HS (55.7%). Participants with HS were found to be significantly younger (P < 0.001) than those without HS, 41.8 (12.2) versus 47.9 (16.0) years old respectively. The BMI was significantly higher in the HS group (P = 0.030). HS patients were relatively more likely than patients without HS to smoke tobacco and less likely to be ex- or non-smokers. However, no significant differences were found in the amount of pack years between currently smoking patients with and without HS (P = 0.549).

The Montreal classification shows that significantly more IBD patients with HS had an early onset of IBD compared with the IBD patients without HS. Additionally, late onset of IBD was more frequent in IBD patients without HS. No differences were found in the localization of IBD between IBD with HS and IBD without HS. Stricturing behaviour of CD was more frequent in IBD without HS

whereas perianal disease occurred more often in CD patients with HS. Disease extent and severity did not differ between UC with HS and UC without HS. IBD patients with HS were significantly more often treated with anti-TNFα therapy and had more intestinal resections.

Table 2 Patient characteristics of IBD without HS versus IBD with HS

	IBD without HS n = 1126 n (%)	IBD with HS n = 134 n (%)	P value
Type IBD			
Crohn's disease	538 (47.8)	96 (71.6)	< 0.001
Ulcerative colitis	588 (52.2)	38 (28.4)	< 0.001
Gender			
Female	627 (55.7)	113 (84.3)	< 0.001
Male	499 (44.3)	21 (15.7)	< 0.001
Smoking status[a]			
Non-smokers	470 (41.9)	46 (34.3)	0.095
Former smokers	452 (40.3)	48 (35.8)	0.005
Current smokers	199 (17.8)	40 (29.9)	0.001
Montreal classification			
Age of onset			
16 years or younger	56 (10.6)	8 (11.4)	0.830
17-40 years	336 (63.5)	55 (78.6)	0.013
Over 40 years	137 (25.9)	70 (10)	0.003
Localization (CD)			
Terminal Ileum (L1)	107 (36.0)	15 (26.3)	0.158
Colon (L2)	56 (18.9)	13 (22.8)	0.490
Ileocolon (L3)	101 (34.0)	22 (38.6)	0.505
Upper gastrointestinal (L4)	5 (1.7)	0 (0.0)	1.000
L1 +L4	11 (3.7)	3 (5.3)	0.479
L2 +L4	7 (2.4)	1 (1.8)	1.000
L3+L4	10 (3.4)	3 (5.3)	0.447
Behaviour (CD)			
Non stricturing, non penetrating (B1)	123 (38.2)	24 (37.5)	0.916
B1 + perianal (P)	32 (9.9)	16 (25.0)	0.001
Stricturing (B2)	80 (24.8)	6 (9.4)	0.007
B2 + perianal (P)	38 (11.8)	8 (12.5)	0.875
Penetrating (B3)	31 (9.6)	5 (7.8)	0.648
B3 + perianal (P)	18 (5.6)	5 (7.8)	0.493
Extent (UC)			
Proctitis (E1)	41 (16.1)	4 (22.2)	0.511
Left-sided UC (E2)	83 (32.5)	4 (22.2)	0.442
Extensive UC (E3)	131 (51.4)	10 (55.6)	0.731
Severity (UC)			
Remission (S0)	10 (4.2)	0 (0.0)	1.000

Table 2 *continued*

	IBD without HS n = 1126 n (%)	IBD with HS n = 134 n (%)	P value
Mild (S1)	72 (30.1)	5 (33.3)	0.793
Moderate (S2)	96 (40.2)	5 (33.3)	0.600
Severe (S3)	61 (25.5)	5 (33.3)	0.503
Immunosuppressive agents			
Yes	259 (44.6)	38 (50.0)	0.372
No	322 (55.4)	38 (50.0)	0.372
Anti-TNFα therapy			
Yes	111 (19.1)	29 (38.2)	< 0.001
No	470 (80.9)	47 (61.8)	< 0.001
Surgical resection			
Yes	267 (39.1)	52 (58.4)	< 0.001
No	416 (60.9)	37 (41.6)	< 0.001
Pack years[a] median (range)			
Current smokers	9.0 (0.03-47.50)	8.0 (0.15-35.00)	0.549
Former smokers	7.4 (0.05-76.00)	8.7 (0.25-80.00)	0.590
Age at onset:			
IBD	-	25.5 (7.00-63.00)	
HS	-	25.0 (6.00-68.00)	
Age[b] (mean years) (SD)	47.9 (16.0)	41.8 (12.2)	< 0.001
BMI (mean kg/m^2)[a] (SD)	25.0 (4.4)	26.1 (5.4)	0.030

[a]Missing values for smoking status (n = 5), pack years (n = 145) and BMI (n = 23), Montreal classification (age of onset n = 661, localization n = 280, behaviour n = 248, extent n = 353, severity n = 372), use of immunosuppressive agents and/or anti-TNFα therapy (n = 603) and surgical resection (n = 488).
[b]Age was determined at date of closing database (May 21st, 2014).
IBD: inflammatory bowel disease; HS: hidradenitis suppurativa.

CD with HS Versus UC with HS

The BMI in CD+HS (25.3 (5.2) kg/m^2) was significantly lower (P = 0.009) than in UC+HS (28.0 (5.5) kg/m^2). The CD+HS population significantly (P = 0.002) tend to smoke more often, whereas the patients in the UC+HS group are more likely to be former smokers (P < 0.001). Patients with CD+HS demonstrated to have developed their intestinal disease significantly earlier (24.0 (9.00-63.00) years) than UC+HS patients (30.0 (7.00-52.00) years) with P = 0.018. No significant differences were found for age, gender, and onset of HS. All the characteristics described above are shown in Table 3.

Parameters Associated with HS in IBD

Multivariable logistic analysis described the relation between HS and its different parameters in an association model, as shown in Table 4. Female gender was the best independent associated parameter for having HS in IBD (OR = 3.494, P < 0.001). In addition, CD patients were more likely than UC patients to have HS (OR = 2.112, P < 0.001). Smoking, a higher BMI, and younger age seem to

contribute in developing HS as well (OR = 1.910, OR = 1.075, and OR = 0.973 respectively). The model including the associated parameters fitted significantly better than without these parameters, with a chi-square value of 94.6 and a P value < 0.001.

Table 3 Patient Characteristics of CD with HS Versus UC with HS

	CD + HS (n = 96) n (%)	UC + HS (n = 38) n (%)	P value
Gender			
Female	84 (87.5)	29 (76.3)	0.108
Male	12 (12.5)	9 (23.7)	0.108
Smoking status[a]			
Non-smokers	33 (34.4)	13 (34.2)	0.986
Former smokers	27 (28.1)	21 (55.3)	< 0.001
Current smokers	36 (37.5)	4 (10.5)	0.002
Pack years[a] median (range)			
Current smokers	8.0 (0.30-35.00)	9.3 (0.15-33.75)	0.825
Former smokers	6.0 (0.38-80.00)	10.6 (0.25-30.00)	0.191
Median age at onset (range)			
IBD	24.0 (9.00-63.00)	30.0 (7.00-52.00)	0.018
HS	26.0 (6.00-68.00)	23.0 (10.00-65.00)	0.277
Age[b] (mean years) (SD)	41.1 (12.6)	43.6 (10.8)	0.277
BMI (mean kg/m^2)[a] (SD)	25.3 (5.2)	28.0 (5.5)	0.009

[a]Missing values for smoking status (n = 5), pack years (n = 145) and BMI (n = 23)
[b]Age was determined at date of closing database (May 21st, 2014).
IBD: inflammatory bowel disease; HS: hidradenitis suppurativa.

Table 4 Multivariate logistic regression of risk factors for hidradenitis suppurativa

	β	P value	Odds Ratio	95% CI
Gender	1.251	< 0.001	3.494	2.138-5.712
IBD type	0.748	< 0.001	2.112	1.389-3.213
Smoking behaviour	0.647	0.010	1.910	1.167-3.126
BMI (kg/m^2)	0.073	< 0.001	1.075	1.035-1.118
Age (years)	-0.028	< 0.001	0.973	0.959-0.987

IBD: inflammatory bowel disease

Genetic Association Analysis

We performed a within cases allelic association analysis for 59 IBD patients with HS and 293 IBD patients without HS. We did not identify any signals at genome wide significance level (P value < 1.0×10^{-8}). Two suggestive genetic association signals were observed on chromosomes 4 and 5 (Figure 1, Table 5). The first signal on chromosome 4, rs2014777 (P value of 7.5×10^{-5}; OR = 2.3) resides in an intergenic region between *SULT1B1* and *SULT1E1*. The genetic association signal on chromosome

5, rs10057395 (P value 7.2 × 10^{-5}; OR = 0.4) is in a genomic region harbouring *ELOVL7*. Both SNPs passed our QC measures (including manual cluster plot inspection) and the regions show additional SNPs showing signals of suggestive evidence for association (Figure 2A, B).

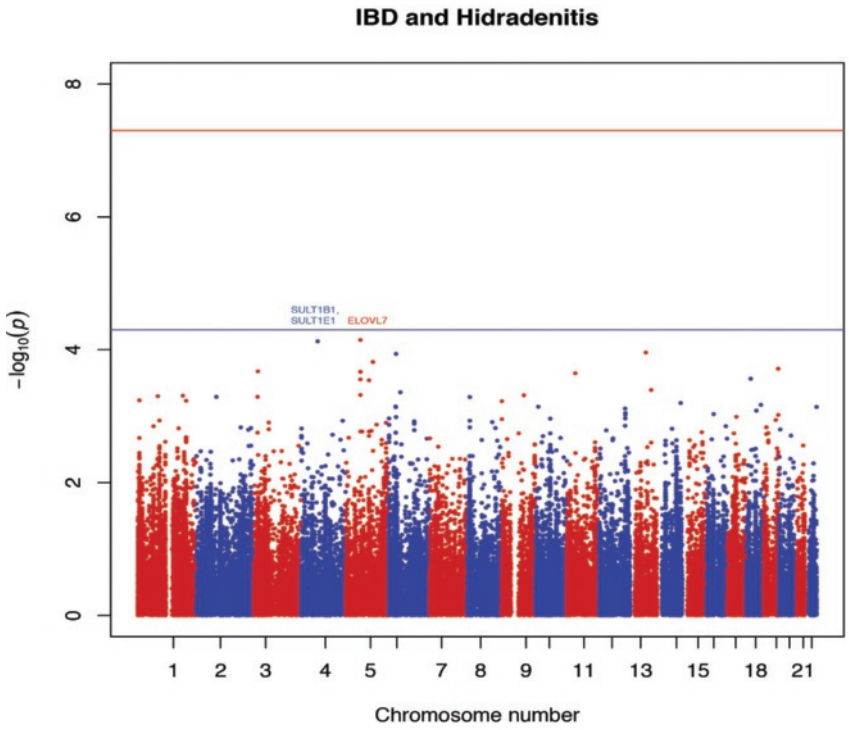


Figure 1 Manhattan plot of association tests of all single nucleotide polymorphism that passed stringent QC; The red horizontal line indicates the threshold for genome-wide significance (P value of 1.0 x 10^{-8}). The blue horizontal line indicates a P value of 1.0 x 10^{-5}.

Table 5 Genes with the highest association for HS in IBD

Chromosome	Position	SNP	Gene	OR	P value
4	70706332	rs2014777	*SULT1B1, SULT1E1*	2.25	7.48 x 10^{-5}
5	60124551	rs10057395	*ELOVL7*	0.43	7.15 x 10^{-5}

Position is relative to human reference genome GRCh37; HS: hidradenitis suppurativa; IBD: inflammatory bowel disease; SNP: single nucleotide polymorphism; OR: odds ratio.

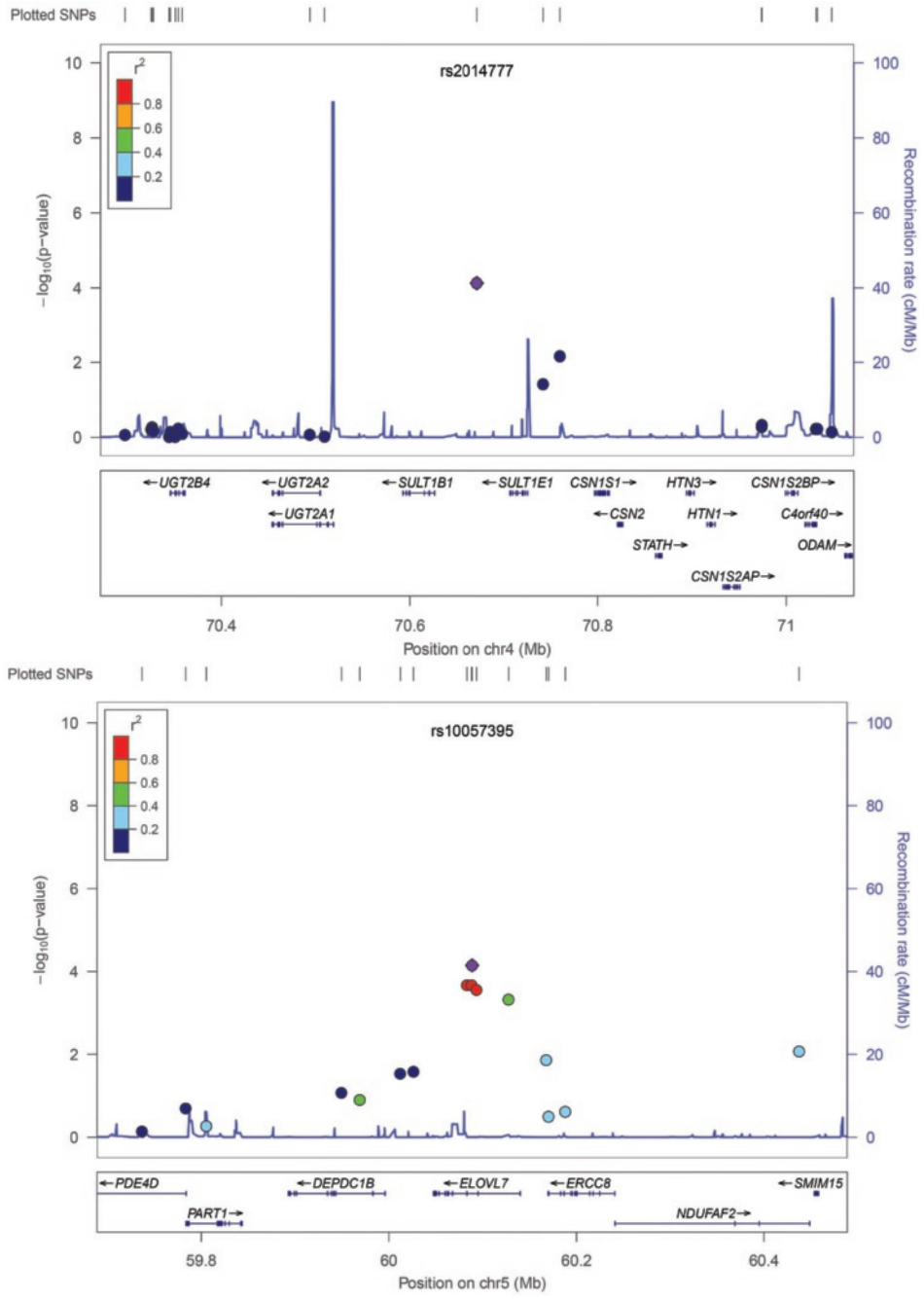


Figure 2 Regional association plots of the 2 suggestive genetic association signals (purple triangles). Filled in circles are genotyped SNP's from the Immunochip. The colour illustrates linkage disequilibrium with the associated SNP (2A: rs2014777 2B: rs10057395).

Discussion

The first goal of this study was to investigate the association between HS and IBD. The prevalence of HS in our IBD population is 6.8% to 10.6%, significantly higher than the prevalence in the general population (1%) confirming the association of the 2 diseases.[2] Moreover, we demonstrate that HS is also associated with CD and UC separately. In 2010 and 2014, van der Zee *et al* found a conspicuous higher prevalence of HS in IBD patients compared with our prevalence (16% and 23%, versus 6.8%-10.6% in our study).[5,24] We believe our data is more accurate because of the verification of HS. Questionnaires used for the investigation of HS prevalence are only validated in the general population.[25] Because in IBD, especially in CD, enterocutaneous fisteling and perianal skin changes frequently occur, patients can confuse this with HS.

The distribution of IBD type within the HS patients was practically similar to the previous study, with 71.6% CD patients in our population and 71.0% in the study of van der Zee *et al*.[24] The higher prevalence of HS in CD suggests that the association between IBD and HS predominantly exists due to CD. This might be explained by the fact that genetic influence is stronger in CD than in UC.[29]

The influence of HS on IBD severity has up until now never been described. Our results indicate that co-morbid HS is associated with an early onset of IBD in which anti-TNFα therapy and surgical resections are often needed. Moreover, we confirmed the finding of frequent perianal IBD in HS patients as previously described by Yadaz *et al*.[30]

When comparing CD+HS with UC+HS, CD+HS patients seem to be more often current smokers. The same smoking status was found in the study of van der Zee *et al*.[24] The difference in mean BMI between CD+HS and UC+HS was higher in our patients, in favour of CD, compared with the previous study, where mean BMI was nearly equal in the 2 IBD types.[24] A high BMI in CD could be an independent risk factor.

Striking differences can be found when comparison is made between HS patients in our IBD population and HS patients in the general population called regular HS (RHS). The RHS group investigated by Sartorius *et al* [31] in a large Swedish population turned out to be heavier (28.3 kg/m^2) than our IBD+HS group (26.1 kg/m^2). The most reasonable explanation is that IBD patients more often suffer from malnutrition secondary to intestinal disease.[32] On the contrary, in HS, overweight is a well-known risk factor and the BMI of patients is predominantly too high, even though dermatologists strictly insist on patients losing weight.[1,31] The prevalence of smoking is also remarkably higher in RHS patients (70.0%) compared with IBD+HS patients (29.9%).[31,33] HS patients with IBD (35.8%) are more likely than those without IBD (15.0%) to have quit smoking. The explanation for the difference in smoking behaviour might be that the majority of IBD+HS patients suffer from CD. Cessation of smoking improves the course in CD patients, contrary to UC patients,

where cessation even may cause flares of the disease.[34] However, in RHS the influence of quitting smoking on disease course is unknown and stopping seems to make no difference.[35]

The parameters associated with HS in our IBD population correspond to the risk factors in RHS: female gender, smoking, and obesity. The prevalence of currently smoking and heavy HS patients is much higher in the general population than in the IBD population, as mentioned above. Additionally, our data shows that younger IBD patients are more likely to have HS than older participants. This corresponds with previous research, demonstrating that HS mainly initiates in the early twenties.[1]

We identified suggestive signals for genetic association for HS in IBD at 2 genomic loci harbouring the genes *ELOVL7*, *SULT1B1*, and *SULT1E1*. *SULT1B1* and *SULT1E1* belong to the sulfotransferase family, encoding enzymes that catalyse the sulphate conjugation of hormones, drugs, neurotransmitters, and xenobiotic compounds.[36] The suggestive signal for *SULT1E1* might have a relevant link with HS for several reasons. First of all, *SULT1E1* encodes for an enzyme which is involved in oestrogen homeostasis.[37] Oestrogens also seem to play a role in HS. Exacerbations of HS occur frequently during relatively hypo-estrogenic states. It is therefore hypothesized that oestrogens exert a protective effect for the disease.[38] Another reason why *SULT1E1* might be linked to HS is, that it is expressed in abdominal subcutaneous tissue in obese males and females (BMI > 30 kg/m^2). Adiposity is an important risk factor for HS.[11] Moreover, Ahima *et al*[37] found oestrogen expression in abdominal adipose tissue in association with expression of TNF-α. TNF-α plays an important role in the disease pathogenesis of CD and HS and both diseases respond well to anti-TNFα therapy. The gene *ELOVL7* encodes a long-chain fatty acid elongase and has no obvious link with HS.[39]

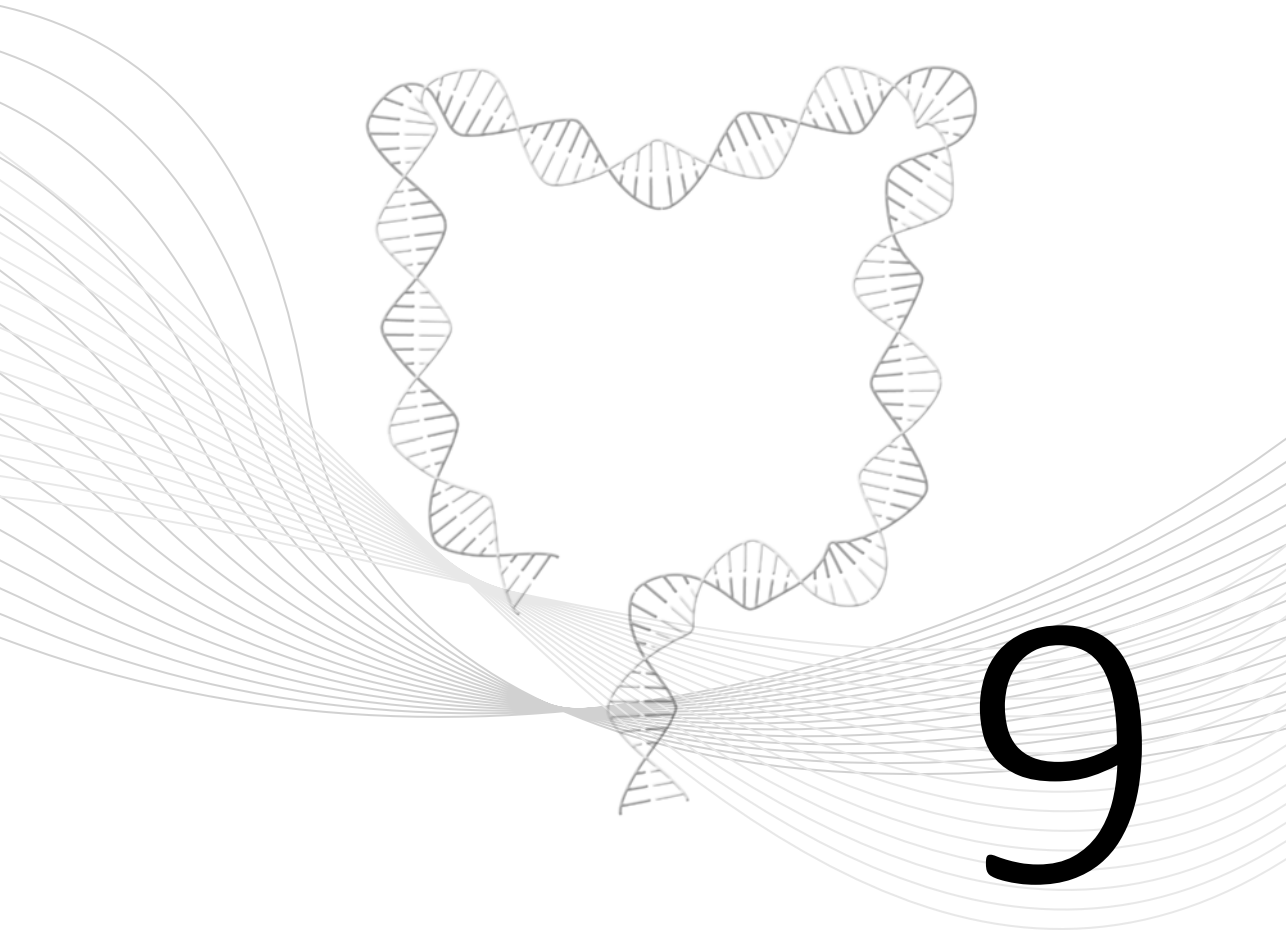
Little is known in literature about the genetic background of HS in IBD patients. Gao *et al* identified a candidate locus for HS at chromosome 1p21.1-1q25.3. However, we could not confirm this association in our IBD cohort.[16] Although the current cohort was underpowered to identify genetic association signals at genome-wide significance level, we observed suggestive evidence for association at 2 genomic loci harbouring potential candidate genes *ELOVL7*, *SULT1B1*, and *SULT1E1*. To confirm these genetic associations for HS in IBD, additional independent cohorts need to be analysed.

In conclusion, this is the first study reporting the association between IBD and HS in such a large cohort. Our study demonstrates the importance of verification of HS in patients recruited from validated questionnaires. In IBD patients, co-morbid HS is associated with an early onset of IBD in which anti-TNFα therapy and surgical resections are often needed. The development of HS in IBD is associated with female gender, CD, smoking, higher BMI, and younger age. Gastroenterologists should therefore pay special attention to this group of patients as they are at risk of developing HS.

References

1. Jemec GB. Clinical practice. Hidradenitis suppurativa. *N Engl J Med* 2012;366:158-164.
2. Revuz JE, Canoui-Poitrine F, Wolkenstein P, *et al*. Prevalence and factors associated with hidradenitis suppurativa: results from two case-control studies. *J Am Acad Dermatol* 2008;59:596-601.
3. Khor B, Gardet A, Xavier RJ. Genetics and pathogenesis of inflammatory bowel disease. *Nature* 2011;474:307-317.
4. Xavier RJ, Podolsky DK. Unravelling the pathogenesis of inflammatory bowel disease. *Nature* 2007;448:427-434.
5. van der Zee HH, van der Woude CJ, Florencia EF, *et al*. Hidradenitis suppurativa and inflammatory bowel disease: are they associated? Results of a pilot study. *Br J Dermatol* 2010;162:195-197.
6. Revuz J. Hidradenitis suppurativa. *J Eur Acad Dermatol Venereol* 2009;23:985-998.
7. van der Zee HH, de Ruiter L, van den Broecke DG, *et al*. Elevated levels of tumour necrosis factor (TNF)-alpha, interleukin (IL)-1beta and IL-10 in hidradenitis suppurativa skin: a rationale for targeting TNF-alpha and IL-1beta. *Br J Dermatol* 2011;164:1292-1298.
8. Giamarellos-Bourboulis EJ, Antonopoulou A, Petropoulou C, *et al*. Altered innate and adaptive immune responses in patients with hidradenitis suppurativa. *Br J Dermatol* 2007;156:51-56.
9. van Wijk F, Cheroutre H. Mucosal T cells in gut homeostasis and inflammation. *Expert Rev Clin Immunol* 2010;6:559-566.
10. Schlapbach C, Hanni T, Yawalkar N, *et al*. Expression of the IL-23/Th17 pathway in lesions of hidradenitis suppurativa. *J Am Acad Dermatol* 2011;65:790-798.
11. Nazary M, van der Zee HH, Prens EP, *et al*. Pathogenesis and pharmacotherapy of Hidradenitis suppurativa. *Eur J Pharmacol* 2011;672:1-8.
12. Rosenfeld G, Bressler B. The truth about cigarette smoking and the risk of inflammatory bowel disease. *Am J Gastroenterol* 2012;107:1407-1408.
13. Higuchi LM, Khalili H, Chan AT, *et al*. A prospective study of cigarette smoking and the risk of inflammatory bowel disease in women. *Am J Gastroenterol* 2012;107:1399-1406.
14. van der Heide F, Dijkstra A, Weersma R, *et al*. Effects of active and passive smoking on disease course of Crohn's disease and ulcerative colitis. *Inflamm Bowel Dis* 2009;15:1199-1207.
15. Chan SS, Luben R, Olsen A, *et al*. Body mass index and the risk for Crohn's disease and ulcerative colitis: data from a European Prospective Cohort Study (The IBD in EPIC Study). *Am J Gastroenterol* 2013;108:575-582.
16. Gao M, Wang PG, Cui Y, *et al*. Inversa acne (hidradenitis suppurativa): a case report and identification of the locus at chromosome 1p21.1-1q25.3. *J Invest Dermatol* 2006;126:1302-1306.
17. Al-Ali FM, Ratnamala U, Mehta TY, *et al*. Hidradenitis suppurativa (or Acne inversa) with autosomal dominant inheritance is not linked to chromosome 1p21.1-1q25.3 region. *Exp Dermatol* 2010;19:851-853.
18. Pink AE, Simpson MA, Brice GW, *et al*. PSENEN and NCSTN mutations in familial hidradenitis suppurativa (Acne Inversa). *J Invest Dermatol* 2011;131:1568-1570.
19. Pan Y, Lin MH, Tian X, *et al*. Gamma-secretase functions through Notch signaling to maintain skin appendages but is not required for their patterning or initial morphogenesis. *Dev Cell* 2004;7:731-743.
20. Blok J, Jonkman M, Horvath B. The possible association of hidradenitis suppurativa and Down syndrome: is increased amyloid precursor protein expression resulting in impaired Notch signalling the missing link? *Br J Dermatol* 2014;170:1375-1377.
21. Savva A, Kanni T, Damoraki G, *et al*. Impact of Toll-like receptor-4 and tumour necrosis factor gene polymorphisms in patients with hidradenitis suppurativa. *Br J Dermatol* 2013;168:311-317.
22. Hugot JP, Laurent-Puig P, Gower-Rousseau C, *et al*. Mapping of a susceptibility locus for Crohn's disease on chromosome 16. *Nature* 1996;379:821-823.
23. Jostins L, Ripke S, Weersma RK, *et al*. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;491:119-124.

24. van der Zee HH, de Winter K, van der Woude CJ, *et al*. The prevalence of hidradenitis suppurativa in 1093 patients with inflammatory bowel disease. *Br J Dermatol* 2014;171:673-675.
25. Esmann S, Dufour DN, Jemec GB. Questionnaire-based diagnosis of hidradenitis suppurativa: specificity, sensitivity and positive predictive value of specific diagnostic questions. *Br J Dermatol* 2010;163:102-106.
26. Spekhorst LM, Visschedijk MC, Alberts R, *et al*. Performance of the Montreal classification for inflammatory bowel diseases. *World J Gastroenterol* 2014;20:15374-15381.
27. Shah TS, Liu JZ, Floyd JA, *et al*. optiCall: a robust genotype-calling algorithm for rare, low-frequency and common variants. *Bioinformatics* 2012;28:1598-1603.
28. Purcell S, Neale B, Todd-Brown K, *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559-575.
29. Halfvarson J, Bodin L, Tysk C, *et al*. Inflammatory bowel disease in a Swedish twin cohort: a long-term follow-up of concordance and clinical characteristics. *Gastroenterology* 2003;124:1767-1773.
30. Yadav S, Varayil JE, Kamal N, *et al*. Hidradenitis suppurativa in patients with inflammatory bowel disease: a population-based study. *Gastroenterology*. [published online ahead of print May 5, 2015).
31. Sartorius K, Emtestam L, Jemec GB, *et al*. Objective scoring of hidradenitis suppurativa reflecting the role of tobacco smoking and obesity. *Br J Dermatol* 2009;161:831-839.
32. Podolsky DK. Inflammatory bowel disease. *N Engl J Med* 2002;347:417-429.
33. van der Heide F, Dijkstra A, Albersnagel F, *et al*. Active and passive smoking behaviour and cessation plans of patients with Crohn's disease and ulcerative colitis. *J Crohns Colitis* 2010;4:125-131.
34. Garcia Rodriguez LA, Gonzalez-Perez A, Johansson S, *et al*. Risk factors for inflammatory bowel disease in the general population. *Aliment Pharmacol Ther* 2005;22:309-315.
35. Simonart T. Hidradenitis suppurativa and smoking. *J Am Acad Dermatol* 2010;62:149-150.
36. Stanley EL, Hume R, Coughtrie MW. Expression profiling of human fetal cytosolic sulfotransferases involved in steroid and thyroid hormone metabolism and in detoxification. *Mol Cell Endocrinol* 2005;240:32-42.
37. Ahima RS, Stanley TL, Khor VK, *et al*. Estrogen sulfotransferase is expressed in subcutaneous adipose tissue of obese humans in association with TNF-alpha and SOCS3. *J Clin Endocrinol Metab* 2011;96:E1153-E1158.
38. Harrison BJ, Read GF, Hughes LE. Endocrine basis for the clinical presentation of hidradenitis suppurativa. *Br J Surg* 1988;75:972-975.
39. Tamura K, Makino A, Hullin-Matsuda F, *et al*. Novel lipogenic enzyme ELOVL7 is involved in prostate cancer growth through saturated long-chain fatty acid metabolism. *Cancer Res* 2009;69:8133-8140.



9

Genomic and expression analyses identify a disease-modifying variant for fibrostenotic Crohn's disease

MC Visschedijk, LM Spekhorst, S-C Cheng, ES van Loo, BHD Jansen, T Blokzijl, H Kil, DJ de Jong, M Pierik, PWJ Maljaars, JC van der Woude, AA van Bodegraven, B Oldenburg, M Löwenberg, VB Nieuwenhuijs, F Imhann, S van Sommeren, R Alberts, RJ Xavier, G Dijkstra, KN Faber, CM Aldaz, RK Weersma, and EAM Festen;
on behalf of the Parelsnoer Institute (PSI) and the Dutch Initiative on Crohn and Colitis (ICC)

J Crohns Colitis 2018 Apr 27;12(5):582-588

Abstract

Background: Crohn's disease (CD) is a chronic inflammatory disease with unpredictable behaviour. More than half of CD patients eventually develop complications such as stenosis, for which they then require endoscopic dilatation or surgery, as no anti-fibrotic drugs are currently available. We aim to identify disease-modifying genes associated with fibrostenotic CD.

Methods: We performed a within-case analysis comparing 'extreme phenotypes' using the Immunochip and replication of the top single nucleotide polymorphisms (SNPs) with Agena Bioscience in two independent case-control cohorts totalling 322 cases with fibrostenotic (recurrent after surgery) and 619 cases with purely inflammatory CD.

Results: Combined meta-analysis resulted in a genome-wide significant signal for SNP rs11861007 ($P = 6.0910^{-11}$), located on chromosome 16, in lncRNA RP11-679B19.1, an lncRNA of unknown function, and close to exon 9 of the *WWOX* gene, which codes for WW domain-containing oxidoreductase. We analysed mRNA expression of *TGF-β* and downstream genes in ileocaecal resection material from ten patients with and without the *WWOX* risk allele. Patients carrying the risk allele (A) showed enhanced colonic expression of *TGF-β* compared to patients homozygous for the wild-type (G) allele ($P = 0.0079$).

Conclusion: We have identified a variant in *WWOX* and in lncRNA RP11-679B19.1 as a disease-modifying genetic variant associated with recurrent fibrostenotic CD and replicated this association in an independent cohort. *WWOX* can potentially play a crucial role in fibrostenosis in CD, being positioned at the crossroads of inflammation and fibrosis.

Introduction

Crohn's disease (CD) is a chronic inflammatory bowel disease (IBD) with unpredictable disease behaviour; more than half of CD patients eventually develop complications such as fistulae and stenosis.[1] Patients with fibrosis-induced stenosis require endoscopic dilatation or surgery, as no anti-fibrotic drugs are currently available. The estimated lifetime risk for surgery in CD patients is 70%, with fibrostenosis as the most common indication for surgery.[2] Within 18 months after surgery, the endoscopic recurrence rate of fibrostenosis is 40-60%.[3]

Fibrostenosis of the intestine is a result of chronic inflammation: the permanent scarring and the consequent luminal narrowing is induced by continuous events of injury and healing. Finally, the intestine is no longer able to restore the normal histological configuration and excessive deposition of extracellular matrix prevails. This leads to thickening of the intestinal wall, which in turn leads to luminal stricturing, finally causing stenosis and obstruction.

Genetic studies of IBD have identified over 240 risk loci associated with disease development.[4,5] A large genotype-phenotype association study was performed, including 34819 IBD patients, which identified three loci (*NOD2*, *MHC* and *MST1*3p21) associated with disease location.[6] However, only a few risk variants are known that influence disease behaviour, such as *SMAD3* with recurrent surgery in CD, *MAGI1* with complicated structuring CD, and *FOXO3*, *XACT* and *IGFBP1-3*with severe disease course in CD.[7-10] Variants associated with disease onset seem not to be associated with disease behaviour, which suggests that the biological pathways that underlie disease behaviour are distinct from those that underlie disease onset.[10] In this study, we aim to identify disease-modifying genes associated with the fibrostenotic phenotype in CD. This will help us gain insight into the biological mechanisms underlying intestinal fibrosis development, which might open new avenues for the treatment of CD.

We performed a within-case analysis comparing 'extreme phenotypes', an approach that provides power to detect associations within disease cohorts that has previously been successfully adopted for other traits, including obesity.[11,12] We used an immune-focused fine mapping chip (Immunochip) with 166,251 single nucleotide polymorphisms (SNPs), in two independent case-control cohorts totalling 322 individuals with multiple (two or more) resections due to confirmed ileal stenosis, which we used as cases, and which we compared to patients with 619 purely inflammatory CD (non-penetrating/non-fibrostenotic with disease duration > 5 years). Following this approach, we identified and replicated a variant within lncRNA RP11-679B19.1 and in *WWOX*, as a disease-modifier, associated with recurrent fibrostenotic CD. We then studied the potential functional consequences of this variant by studying resection material of CD patients carrying the risk allele.

Methods

Subjects

The discovery cohort consisted of 521 Dutch CD patients, selected out of ~1500 Dutch CD patients. Selection was based on reviewing the medical records of the total cohort of 1500 patients. We selected patients on extreme ends of the fibrostenotic phenotype: a case group and a control group. The case group consisted of 242 patients with 'recurrent fibrostenotic' CD, which was defined as the need for multiple (two or more) resections due to confirmed ileal stenosis. The control group consisted of 279 patients with purely inflammatory disease behaviour CD (Montreal classification B1[13]) defined as non-penetrating, non-fibrostenotic CD with a disease duration of > 5 years, without the need for any kind of surgery. All patients of the discovery cohort were also included in the IIBDGC core-phenotype study.[6]

For the replication phase, patients were selected from the Parelsnoer Institute database (http://www.parelsnoer.org/). This database consists of 3394 IBD patients (CD: *n* = 2118) collected from the IBD Centres in all eight University Medical Centres in the Netherlands. Patients already included in the discovery phase were excluded. Eighty CD patients with recurrent fibrostenotic disease and 340 patients with purely inflammatory disease were selected with similar criteria. While the phenotypes of the discovery cohort could be checked against the patients' medical charts, regulations prevented us from doing the same with the replication cohort. Hence, the replication cohort might have some non-cases and/or non-controls admixed. All patients were of Central European descent.

Discovery phase

Immunochip genotyping

In the discovery phase of the study we used genotype data of ~1500 CD patients described by Jostins *et al.*[14] In brief, DNA was extracted from whole blood. An Illumina Immunochip was used to genotype the DNA samples. The Immunochip is a custom Illumina Infinium immune-focused fine-mapping chip comprising 196,524 SNPs selected primarily based on genome wide association study (GWAS) analysis of 12 immune-mediated diseases.

Quality control

Quality control (QC) was performed using PLINKv1.07 software.[15] QC removed SNPs with a minor allele frequency (MAF) less than 0.001, as well as SNPs with more than 10% missing genotypes. In total, 547 patients were selected (248 fibrostenotic CD patients and 299 CD patients with purely inflammatory disease behaviour). After removing duplicate individuals (*n* = 6) and individuals

with more than 15% missing genotype data (*n* = 20), 521 individuals remained (no differential missingness, χ^2 P = 0.67). After QC, the dataset consisted of 521 individuals and 166,251 SNPs with a genotyping call rate of 99%.

Statistical analysis and prioritization of SNPs for replication

After allele association analysis (χ^2 test) and cluster plot inspection, 34 SNPs with a P value < 1.0×10^{-3} remained. To adequately test each locus for association in the replication phase we selected two SNPs per locus where possible. In addition, we decided to test the SNP that previously showed the strongest association with severe CD disease course (rs12212067, *FOXO3*), although this SNP was not associated with fibrostenotic disease in our discovery cohort.[10,16] In total, 43 SNPs passed the design of two plexes with the Assay Design suite of Agena Bioscience (https://seqpws1.agenacx.com/AssayDesignerSuite.html).

Replication phase

Genotyping

Forty-three SNPs were genotyped using the Agena Bioscience Massarray in the replication phase (http://agenabio.com). During QC we excluded SNPs with a Hardy-Weinberg equilibrium P value < 0.0001 in controls and an overall call rate < 90% and individuals with < 85% of SNPs confidently genotyped. After QC, the replication cohort consisted of 80 CD cases with recurrent fibrostenotic disease behaviour, 340 CD patients with purely inflammatory disease and 39 SNPs with a genotype call rate of 99%. Since rs11861007 failed the Agena Bioscience design and it was the only Immunochip SNP in the *WWOX* region, genotyping of SNP rs11861007 was performed using Taqman technology (Applied Biosystems).

Statistical analysis

Allelic association analysis (χ^2 test) and meta-analysis of the complete cohort was performed with PLINKv1.07 software.[15] To test if disease localization could be a cofounding factor, we additionally performed an allelic association analysis for rs11861007 (χ^2 test) in ileal disease patients (*n* = 86) vs colonic disease patients (*n* = 36). Similarly, allelic association analysis for rs11861007 (χ^2 test) in Montreal Class B2 phenotype patients (*n* = 165) vs Montreal Class B1 and B3 phenotype patients (776) was performed to determine whether rs11861007 was associated with the Montreal B2 phenotype (and not recurrent fibrostenotic disease). Logistic regression analyses with disease as a covariate was performed with R statistics.

Given that the prevalence of recurrent stenotic disease in CD was 0.04 in our replication cohort, that the allele frequency of our rs11861007 risk allele A is 0.06 in Caucasian patients of Western and Central European descent (CEU), and that we expected the risk variant to have an odds ratio (OR) of 4.3, we expected to have > 80% power to discover this association in our replication cohort. Ultimately, we only had 60% power to detect the association that we did detect for rs11861007 in our replication cohort at an OR of 2.1 (http://zzz.bwh.harvard.edu/gpc/cc2.html).

SNP annotation

LocusZoom was used to construct regional association plots[17] (**Supplementary Figure S1**). Exploration of the linkage disequilibrium (r^2 > 0.8) was performed with Haploview.[18] We assessed whether associated SNPs had known functional consequences or regulatory features. The Encyclopedia of DNA Elements (ENCODE)[19] was searched using the UCSC Genome Browser.[20] Specifically, SNPs located in the following regulatory features were searched: DNaseI-hypersensitivity sites, transcription factor binding sites, histone modification and DNA-polymerase sites. We tested whether associated variants (or SNPs in high linkage disequilibrium (r^2 > 0.8)) showed an effect on gene expression levels of genes. Expression quantitative trait loci (eQTL) analysis was performed with the eQTL browser (http://genenetwork.nl/bloodeqtlbrowser/), based on non-transformed peripheral blood in 5311 individuals.[21] Additional enhancer analyses were performed with the Fantom5 enhancer atlas.[22]

In silico analysis (rs11861007)

The strongest association for fibrostenotic disease was found in the RP11-679B19.1 lncRNA, and in the *WWOX* gene. The function of the RP11-679B19.1 lncRNA is unknown. To predict in which biological or cellular process and molecular function the *WWOX* gene is involved, we used an in-house developed RNA network tool (http://www.genenetwork.nl).[23] The RNA network uses a method, based on principal component analysis, to build transcriptional profiles for biological pathways, which can be used to predict gene functions. A more detailed description of the method of the RNA network can be found in the paper by Fehrmann *et al*.[23]

Expression analysis

Resection material of CD patients

Subjects

The resection material was collected during surgical resection procedures (samples of the small bowel and colon) from patients for whom we had extensive phenotype data (University Medical Center Groningen). Twenty-nine patients with fibrostenotic CD behaviour were included. The resection material of CD patients was preserved (storage at −80°C) in three parts: the ileum (proximal of the stenosis), the stenosis itself (medial part) and the colon (distal of the stenosis). As stenotic tissue is a final stage in the process of fibrosis formation, we considered the stenotic tissue, medial part of the resection, as not representative. Therefore, we included only samples of the ileum and colon, respectively proximal and distal from the stenosis. Pathology records of the resection material were checked and revealed no severe inflammation or fibrosis at the end of the resection sides, proximal and distal, of the stenosis.

Genotyping and patient selection

After DNA isolation, genotyping of the *WWOX* SNP (rs11861007) in 71 patients was performed using Taqman technology with a call rate of 96% (Life Technologies). We selected the five CD patients who carried the *WWOX*-genotype (risk allele, A) as cases (we found no patients homozygous for the risk allele in our cohort). We selected five matched CD patients with the GG genotype (wild-type) as controls (matched based on age, disease duration, inflammation and stenosis).

Quantitative polymerase chain reaction (qPCR) was performed of *WWOX*, *TGF-β, iNOS*, *IL1-B*, *TNF-α*, *FOXP3*, *α-SMA*, *Collagen Type 1* and downstream genes *PAI-I (SERPINE)* and *CTGF* of the ileocolonic tissue of five CD patients carrying the *WWOX* risk allele and five CD patients homozygous for the *WWOX* wild-type allele (see Online Methods).

Results

In this study we included two independent case-control cohorts totalling 322 recurrent fibrostenotic and 619 purely inflammatory CD cases; 60% of the cohort were female with a mean age at diagnosis of 25 years. The cohorts were selected based on 'extreme phenotypes', resulting in a statistically significant difference in disease location and behaviour between patients (Table 1). After analysis of 166,251 SNPs in 242 fibrostenotic and 279 purely inflammatory CD cases and replication of 34 selected SNPs in an independent cohort of 80 fibrostenotic and 340 purely inflammatory CD

cases, the combined meta-analysis resulted in a genome-wide significant signal for SNP rs11861007 ($P = 6.09 \times 10^{-11}$, OR = 3.2, heterogeneity (I^2) < 75%) (Table 2, **Supplementary Table S1**). The minor heterogeneity for SNP rs11861007 between the discovery and the replication cohort might be caused by a slight admixture of non-cases and/or non-controls in the replication cohort, as described in the Methods section. To assess whether the association between rs11861007 and fibrostenotic disease was not mainly due to ileal disease localization we tested this marker for association with ileal disease localization. We found no association between rs11861007 and ileal disease localization (P = 0.27). Additional logistic regression analyses for the association between SNP rs11861007 and re-fibrosis, with disease location as a covariate, still showed a significant association (P = 0.004). We also tested for an association between rs11861007 and the Montreal Class B2 phenotype (any stricturing disease) in our cohort, but did not find an association (P = 0.25).

Table 1 Clinical characteristics of patients in the combined cohort

	Fibrostenotic CD	Purely inflammatory CD	P value
n (%)	322 (100%)	619 (100%)	
Patient characteristics			
Female, n (%)	180 (56%)	371 (60%)	0.23
Median age of onset, years (IQR 25-75)	24 (18-30)	26 (19-36)	0.38
Disease Location, n (%)			
Montreal-L1, ileal	86 (27%)	93 (15%)	0.06
Montreal-L2, colonic	36 (11%)	291 (47%)	< 0.0001
Montreal-L3, ileocolonic	200 (62%)	235 (38%)	0.0011
Montreal-L4, additional upper disease localization	32 (10%)	55 (9%)	1
Disease Behaviour, n (%)			
Montreal-B1, inflammatory	0	619 (100%)	
Montreal-B2, stricturing	165 (51%)	0	
Montreal-B3, penetrating	157 (49%)	0	
Time until surgery			
Disease duration in years from diagnosis until first surgery, mean (SD)	4.1 (6.5)	NA	

This table provides the clinical characteristics of patients with recurrent fibrostenotic and purely inflammatory CD in the combined cohort. Disease location and behaviour is based on the Montreal classification for CD. Chi-squared and Mann-Whitney U-tests (only for age of onset) are used to calculate the P value.
CD: Crohn's disease; IQR: interquartile range.

Table 2 Allelic association analyses

					Discovery cohort				Replication cohort				Meta-analysis	
Chr.	SNP	Position (Hg19)	Candidate Gene	Risk Allele	Risk allele Frequency in cases	Risk allele Frequency in controls	P value	Odds Ratio	Risk allele Frequency in cases	Risk allele Frequency in controls	P value	Odds Ratio	P value	Odds Ratio
16	rs11861007	79238685	*WWOX*	A	0.19	0.05	1.26×10^{-11}	4.3	0.13	0.07	0.01	2.1	6.09×10^{-11}	3.2
22	rs371513	21988599	*CCDC116*	A	0.22	0.13	1.02×10^{-4}	1.9	0.20	0.15	1.47×10^{-1}	1.4	7.97×10^{-5}	1.7
5	rs55965691	593812	*CEP72*	A	0.008	0.06	1.60×10^{-5}	0.14	0.025	0.06	5.97×10^{-2}	0.38	9.37×10^{-5}	0.23
5	rs6883704	599390	*CEP72*	C	0.01	0.06	5.04×10^{-5}	0.18	0.025	0.06	5.97×10^{-2}	0.38	1.01×10^{-4}	0.25
3	rs17033143	10630258	*ATP2B2*	A	0.14	0.07	1.84×10^{-4}	2.2	0.09	0.07	3.24×10^{-1}	1.36	3.51×10^{-4}	1.9

The five most significant allelic association analysis results (χ^2 test) for the 166,251 SNPs in the discovery cohort (242 patients with recurrent fibrostenotic CD vs 279 patients with purely inflammatory CD) are presented. Replication was performed for 39 SNPs in a replication cohort (80 patients with recurrent fibrostenotic CD vs 340 patients with purely inflammatory CD). Meta-analysis was performed in the combined cohort.
CD: Crohn's disease; SNP: single nucleotide polymorphism; OR: odds ratio.

SNP annotation and in silico analyses

rs11861007 is located on chromosome 16, in an intron close to exon 9 of the *WWOX* gene, which codes for WW domain-containing oxidoreductase. Neither rs11861007 nor SNPs in high linkage disequilibrium ($r^2 > 0.8$) have known functional or regulatory features (eQTL or regulatory elements assessed in the ENCODE Encyclopedia of DNA Elements or FANTOM5 enhancers).[22] *WWOX* is the only coding gene in the locus (defined as 250 kb on either side of our top hit) making it the most likely positional candidate gene. The SNP is also located in the lncRNA RP11-679B19.1, in which it might affect folding, but for this lnRNA there is also no eQTL effect, and no function is known for RP11-679B19.1.

Co-transcriptional pathway analysis using an in-house-developed RNA network tool[23] predicts *WWOX* involvement with cellular components in the *extracellular matrix compartment* ($P = 1.41 \times 10^{-3}$) and an association with *collagen binding* ($P = 9.78 \times 10^{-5}$). See **Supplementary Table S2** for involvement in other pathways.

Expression analyses

Normal ileal and colonic tissue residing proximal to and distal from the stenotic part was sampled from five CD patients heterozygous for the *WWOX* risk allele (A) and five CD patients homozygous for the *WWOX* wild-type allele (G). *WWOX* mRNA levels were similar in both the ileal and the colonic tissue of patients with or without the risk allele (**Supplementary Figure S2**). In contrast, *TGF-β* expression was significantly higher in the colonic tissue of risk-allele-carrying patients compared to patients homozygous for the wild-type allele (Mann-Whitney U-test, P = 0.0079) (Figure 1A). Similarly, downstream targets involved in fibrosis, such as Connective Tissue Growth Factor (*CTGF*) and matricellular *PAI-1* (*SERPINE1*), showed a trend towards increased expression in the risk-allele-carrying CD patients compared to the patients homozygous for the wild-type allele (Figure 1B; see **Supplementary Figure S2** for all qPCR results).

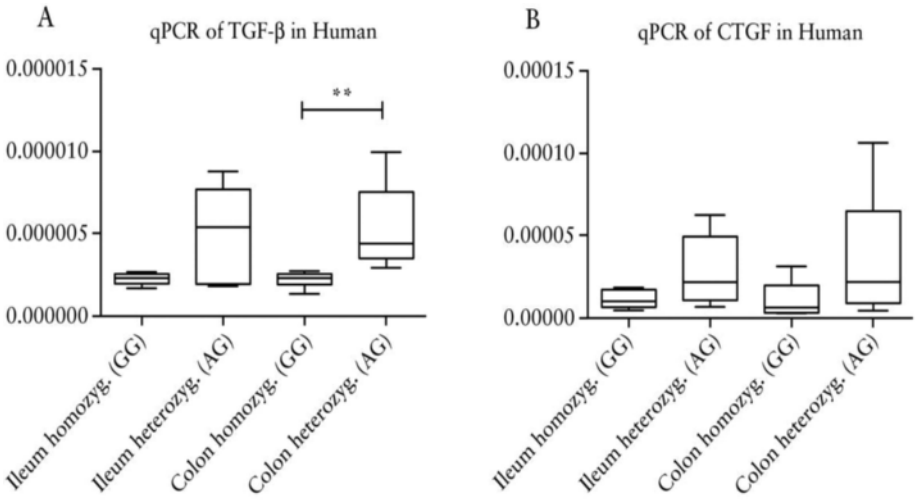


Figure 1 TGF-β + CTGF expression in human and macrophage polarization of *WWOX*. TGF-β is a crucial factor in the equilibrium between inflammation and fibrosis. (A, B) *TGF-β* and *CTGF* (profibrotic downstream gene of TGF-β) expression in the non-stenotic ileocaecal resection tissue from five CD patients carrying the risk allele (AG) and five CD patients homozygous for the wild-type allele (GG).
CD: Crohn's disease; qPCR: quantification polymerase chain reaction; TGF-β: transforming growth factor; CTGF: connective tissue growth factor; homozyg: homozygous; heterozyg: heterozygous.

Discussion

In this study, we have identified a variant in the *WWOX* gene as a disease-modifier associated with a recurrent fibrostenotic phenotype in CD, and we have replicated this association in an independent cohort. We found that patients carrying the risk allele show enhanced profibrotic signalling, through higher expression of tumour growth factor-β (TGF-β), in the colon. This suggests that the genetic variant in *WWOX* is associated with a decrease of WWOX function. Alternatively, the effect could be caused by a configuration change of the lnRNA RP11-679B19.1, which has recently been recorded in the same locus and is, yet, of unknown function.

The genetic variants that contribute to disease behaviour can be different from the variants that contribute to disease susceptibility. This has been shown for a variant affecting *FOXO3A* expression, which was found to be significantly associated with disease prognosis but which was not associated with CD susceptibility.[16] In this study, we confirm the concept that disease-modifying genes can differ from variants contributing to disease development. The SNP found to be most closely associated (rs11861007) has not been associated with disease development in previous studies.[4] Previous genotype-phenotype studies in IBD patients have not studied the specific phenotype in the present study: recurring ileal fibrostenosis in CD.[6] Previous studies focusing on genetic variants associated

with a single fibrotic event in CD found associations with variants in *SMAD3* and *MAGI*, but due to the diverse coverage of the genotyping platforms we could not replicate these findings.[7,8] We could also not confirm a previously described association between *NOD2* variants and ileal fibrostenotic disease, which might be due to the fact that we studied recurrent fibrostenosis, not necessarily ileal, or because the *NOD2* variants are relatively rare in our cohort.

The SNP (rs11861007) is located in an intron close to exon 9 of the *WWOX* gene. Although we could not find an eQTL effect, regulatory features or enhancer activity, the *WWOX* gene is the only coding gene in the locus (defined as 250 kb on either side of our top hit) making it the most likely positional candidate gene. The SNP is also located in the lncRNA RP11-679B19.1, in which it might affect folding, but for this lnRNA there is also no eQTL effect, and no function is known for RP11-679B19.1. *WWOX* is a known tumour suppressor gene and encodes a protein that contains two WW domains and a short-chain dehydrogenase/reductase domain (SRD). The mechanism of tumour suppression of *WWOX* involves apoptosis, modulation of the extracellular matrix and modulation of cell bioenergetics[24] Genome-wide association studies have shown that *WWOX* also plays a role in the pathogenesis of pulmonary fibrosis.[25] Targeted deletion of *Wwox* in epithelium in the mammary gland increased fibronectin levels[24] and conditional deletion of *Wwox* in the mammary gland significantly upregulated multiple collagen genes.[26] Moreover, molecular functions predicted by our RNA network tool show that *WWOX* plays a role in fibrosis formation, in agreement with the previously described literature reports.

The main mediator between intestinal inflammation and fibrosis in IBD is TGF-β,[27] which is overexpressed in intestinal tissue in CD patients.[28] Upon TGF-β stimulation, *WWOX* acts as an inhibitor of SMAD3 transcriptional activity by sequestering it in the cytoplasm.[29] We show an enhanced *TGF-β* expression in CD patients carrying the *WWOX* risk allele. TGF-β stimulates downstream signalling pathways resulting in expression of several profibrotic genes, including *CTGF*. qPCR analysis of *CTGF* in this study showed a trend towards increased expression in the *WWOX* risk-allele-carrying CD patients compared to the patients homozygous for the *WWOX* wild-type allele. We conclude that *WWOX* risk-allele-carrying individuals have enhanced TGF-β expression with a trend towards elevated expression of profibrotic genes as a downstream effect.

There are some limitations to this study. First, the cohort size is relatively small. However, by using a within-case analysis comparing 'extreme phenotypes' we increased the power of our analysis. This approach provides power to detect associations within disease cohorts and the approach has been successfully adopted for other traits.[11,12] Moreover, we replicated our findings in an independent cohort. Second, the upregulation of TGF-β in the ileal part of the resection material (upstream of the stenosis) in patients carrying the *WWOX* risk allele was a trend and not statistically significant. However, in the colonic part of the resection material (downstream of the stenosis) in patients carrying the *WWOX* risk allele we do show statistically significant

enhanced *TGF-β* expression. All available *WWOX* risk-allele-carrying patients in our centre were included, but because the risk allele has a low frequency this number is quite small, which means our study is relatively underpowered. Increasing the sample size may turn the trend we observe into a significant association. Finally, the WWOX–SMAD3–TGF-β pathway has been described previously[29] and it has been proven that the proteins encoded by the genes interact, although the exact pathways through which they interact have not yet been elucidated, making it difficult to interpret the results from our study.

In conclusion, we have identified and replicated *WWOX* as a disease-modifying gene associated with the recurrent fibrostenotic phenotype in CD. Our expression analyses suggest a functional effect of the risk allele through enhanced expression of *TGF-β* in risk-allele carriers, indicating profibrotic expression. CD patients carrying the *WWOX* risk allele appear to have a profibrotic profile. To avoid fibrotic complications, it might be advisable to refrain from prescribing anti-inflammatory medication that enhances TGF-β signalling in intestinal fibroblasts in these patients.

Supplementary Data

Supplementary data for this article can be found at *Journal of Crohn's and Colitis* Online.

References

1. Cosnes J, Cattan S, Blain A, *et al*. Long-term evolution of disease behaviour of Crohn's disease. *Inflamm Bowel Dis* 2002;8:244-50.
2. Munkholm P, Langholz E, Davidsen M, *et al.* Disease activity courses in a regional cohort of Crohn's disease patients. *Scand J Gastroenterol* 1995;30:699-706.
3. De Cruz P, Kamm MA, Hamilton AL, *et al.* Crohn's disease management after intestinal resection: a randomised trial. *Lancet* 2014;6736:1-11.
4. Liu JZ, van Sommeren S, Huang H, *et al.*; International Multiple Sclerosis Genetics Consortium; International IBD Genetics Consortium. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. *Nat Genet* 2015;47:979-86.
5. de Lange KM, Moutsianas L, Lee JC, *et al*. Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat Genet* 2017;49:256-61.
6. Cleynen I, Boucher G, Jostins L, *et al.*; International Inflammatory Bowel Disease Genetics Consortium. Inherited determinants of Crohn's disease and ulcerative colitis phenotypes: a genetic association study. *Lancet* 2016;387:156-67.
7. Fowler SA, Ananthakrishnan AN, Gardet A, *et al*. SMAD3 gene variant is a risk factor for recurrent surgery in patients with Crohn's disease. *J Crohns Colitis* 2014;8:845-51.
8. Alonso A, Domènech E, Julià A, *et al*. Identification of risk loci for Crohn's disease phenotypes using a genome-wide association study. *Gastroenterology* 2015;148:794-805.
9. Ananthakrishnan AN, Xavier RJ. How does genotype influence disease phenotype in inflammatory bowel disease? *Inflamm Bowel Dis* 2013;19:2021-30.
10. Lee JC, Biasci D, Roberts R, *et al.*; UK IBD Genetics Consortium. Genome-wide association study identifies distinct genetic contributions to prognosis and susceptibility in Crohn's disease. *Nat Genet* 2017;49:262-8.
11. Plomin R, Haworth CM, Davis OS. Common disorders are quantitative traits. *Nat Rev Genet* 2009;10:872-8.
12. Wang K, Li WD, Zhang CK, *et al*. A genome-wide association study on obesity and obesity-related traits. *PLoS ONE* 2011;6:e18939.
13. Silverberg MS, Satsangi J, Ahmad T, *et al.* Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. *Can J Gastroenterol* 2005;19(Suppl A):5A-36A.
14. Jostins L, Ripke S, Weersma RK, *et al.*; International IBD Genetics Consortium (IIBDGC). Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;491:119-24.
15. Purcell S, Neale B, Todd-Brown K, *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007;81:559-75.
16. Lee JC, Espéli M, Anderson CA, *et al.*; UK IBD Genetics Consortium. Human SNP links differential outcomes in inflammatory and infectious disease to a FOXO3-regulated pathway. *Cell* 2013;155:57-69.
17. Pruim RJ, Welch RP, Sanna S, *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* 2010;26:2336-7.
18. Barrett JC, Fry B, Maller J, *et al.* Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 2005;21:263-5.
19. ENCODE Project Consortium. The ENCODE (ENCyclopedia Of DNA Elements) project. *Science* 2004;306:636-40.
20. Rosenbloom KR, Dreszer TR, Pheasant M, *et al.* ENCODE whole-genome data in the UCSC Genome Browser. *Nucleic Acids Res* 2010;38:D620-5.
21. Westra HJ, Peters MJ, Esko T, *et al*. Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nat Genet* 2013;45:1238-43.
22. Andersson R, Gebhard C, Miguel-Escalada I, *et al.* An atlas of active enhancers across human cell types and tissues. *Nature* 2014;507:455-61.

23. Fehrmann RS, Jansen RC, Veldink JH, *et al*. Trans-eQTLs reveal that independent genetic variants associated with a complex phenotype converge on intermediate genes, with a major role for the HLA. *PLoS Genet* 2011;7:e1002197.

24. Abdeen SK, Salah Z, Khawaled S, *et al*. Characterization of WWOX inactivation in murine mammary gland development. *J Cell Physiol* 2013;228:1391-6.

25. Loth DW, Soler Artigas M, Gharib SA, *et al*. Genome-wide association analysis identifies six new loci associated with forced vital capacity. *Nat Genet* 2014;46:669-77.

26. Ferguson BW, Gao X, Kil H, *et al*. Conditional Wwox deletion in mouse mammary gland by means of two Cre recombinase approaches. *PLoS One* 2012;7:e36618.

27. Rieder F, Fiocchi C. Intestinal fibrosis in inflammatory bowel disease - Current knowledge and future perspectives. *J Crohns Colitis* 2008;2:279-90.

28. Burke JP, Ferrante M, Dejaegher K, *et al*. Transcriptomic analysis of intestinal fibrosis-associated gene expression in response to medical therapy in Crohn's disease. *Inflamm Bowel Dis* 2008;14:1197-204.

29. Ferguson BW, Gao X, Zelazowski MJ, *et al*. The cancer gene WWOX behaves as an inhibitor of SMAD3 transcriptional activity via direct binding. *BMC Cancer* 2013;13:593.

10

Immunogenicity to anti-TNFα: a non-HLA association

LM Spekhorst, DJ Buurman, T Billiet, A Vich Villa, EAM Festen, G Dijkstra*,
S Vermeire*, RK Weersma*
*Authors contributed equally

Manuscript in preparation

Abstract

Introduction: Some patients treated with anti-TNFα develop anti-drug antibodies (ADAs), and these can result in a loss of response. Because immunogenicity to anti-TNFα plays a role in the loss of response to anti-TNFα therapy, our aim was to replicate known HLA regions and identify novel (non-HLA) genetic regions associated with the development of ADAs in patients with inflammatory bowel disease (IBD).

Methods: We used two independent cohorts. The UMCG cohort (Netherlands) consisted of 61 IBD patients with ADAs and 163 patients without ADAs. The Leuven cohort (Belgium) consisted of 77 IBD patients with ADAs and 115 patients without ADAs. ADAs were measured using two different in-house-developed ELISA assays. Imputation of the HLA-region was performed with the HIBAG package. Immunochip genome-wide imputation was done using the Michigan imputation server. For the association analysis a cut-off P value of < 0.05 was used. For the genome-wide association meta-analysis a P value of $< 1.0 \times 10^{-8}$ was used.

Results: The HLA-DQA1*05 allele was replicated ($P = 1.0 \times 10^{-3}$) in 121 IBD patients positive for ADAs to anti-TNFα and 239 IBD patients negative for ADAs. We were not able to replicate a previously documented association with the HLA-DRB1*03 allele (P = 0.40). Our genome-wide association meta-analysis for 121 IBD patients with ADAs to anti-TNFα and 239 IBD patients negative for ADAs showed eight suggestive association signals (P values $< 10 \times 10^{-5}$).

Conclusion: We replicated the HLA-DQA1*05 allele associated with anti-TNFα ADAs and identified eight suggestive association signals in non-HLA regions that need to be replicated in larger cohorts. Our results suggest that immunogenicity studies might be useful as biomarkers to predict individual treatment response to anti-TNFα therapy.

Introduction

Inflammatory Bowel Disease (IBD) is a chronic inflammatory disease of the gastrointestinal tract. Patients with a severe disease phenotype—one which is unresponsive to immunomodulators—often require therapy with so-called biologicals: drug products derived from biological sources. Anti-tumour necrosis factor alpha (anti-TNFα) agents such as infliximab and adalimumab are highly effective in most patients, but they are expensive and their side-effects can be severe. Furthermore, approximately 25% of patients do not show response after induction therapy (primary non-responders) and 23-46% of patients, after initially responding to induction therapy, experience loss of response within one year (secondary non-responders).[1] While there are several causes of this loss of response to anti-TNFα treatment, the primary cause is formation of anti-drug antibodies (ADAs).[2] Strategies to prevent ADAs are scheduled dosing to maintain stable trough drug levels and co-administration of an immunomodulator (e.g. thiopurines).[3] However, we are currently unable to predict or differentiate which patients are at risk of developing ADAs to anti-TNFα.

As the adaptive immune system plays an important role in the development of ADAs, it is feasible that genetic factors play a role in this immunogenicity; however data on genetic associations with ADA development are scarce. A large study from the United Kingdom assessed immunogenicity in 1284 CD patients who had been using anti-TNFα for at least 12 months. They identified a genome-wide association signal on chromosome 6 and demonstrated that this signal was driven by HLA-DQA1*05.[4] Another study from Leuven, Belgium assessed immunogenicity in 76 IBD patients who developed antibodies to infliximab and compared them to 116 matched IBD controls who did not develop antibodies during at least two years of infliximab maintenance therapy. These authors showed that HLA-DRB1*03 was associated with ADAs in IBD patients treated with infliximab.[5] Interestingly the HLA-DQA1*05 allele is correlated with the HLA-DRB1*03 allele. The HLA-DQA1*05 signal could thus be driving the HLA-DRB1*03 association signal. This highlights the need for additional studies to increase our power to discover associated variants. At the same time, we need better predictors of response to anti-TNFα therapy in order to enable tailored therapy for individual patients, which would both improve patient care and lower costs. The aim of this study was thus twofold: 1) to identify and confirm HLA regions (HLA-DQA1*05 and HLA-DRB1*03) associated with development of ADAs to anti-TNFα in IBD patients, and 2) to identify novel (non-HLA) genetic regions associated with the development of ADAs.

Methods

Cohort description and sample selection

In this study we used two independent cohorts. The first was collected at the University Medical Center Groningen (UMCG), the Netherlands, and consisted of 61 IBD patients positive for ADAs to infliximab or adalimumab and 163 patients negative for ADAs. ADAs were measured at one time-point during treatment with an in-house-developed ELISA (Sanquin, Amsterdam).

The second cohort was collected at the University Hospitals of Leuven, Belgium and consisted of 77 IBD patients positive for ADAs to infliximab and 115 patients negative for ADAs. Patients positive for ADAs developed ADAs during infliximab treatment. Patients negative for ADAs had been using infliximab for a minimum of two years and had not developed ADAs. ADAs were measured with an in-house-developed ELISA (KU, Leuven).

Genotyping

All patients were genotyped using the Immunochip: a custom-made genotyping array containing ~200,000 genetic variants for 12 auto-immune mediated diseases derived from genome-wide association analysis.[6]

Genotype calling and quality control

Genotype calling was performed per batch separately with Opticall using default settings.[7] Quality control was performed using PLINK.[8] Samples with a call rate ≤ 0.99 were excluded from the datasets. Variants with a call rate ≤ 0.99 or a minor allele frequency (MAF) ≤ 0.001 and variants with a Hardy-Weinberg equilibrium P value ≤ 0.0001 were also excluded. To identify ancestry outliers, samples were merged with 1000 Genomes Project data and outliers identified using multidimensional scaling plots. These ancestry outliers were then removed from the dataset. To prepare the data for imputation, we used GenotypeHarmonizer[9] and VCFtools.[10] Strands were aligned and identifiers updated to the 1000 Genomes data in genome build GRCh37. The datasets were then combined. Duplicates and related samples were identified and removed.

HLA imputation and association analysis

Imputation of the HLA-region was performed using the HIBAG package[11] with Illumina ImmunoChip hg19 (2-digit resolution) as a reference panel. Association analysis was performed by ANOVA test using an additive model. A P value < 0.05 was considered to be statistically significant.

For the replication of HLA-DQA1*05 both cohorts (UMCG and Leuven) were combined, with a total of 121 IBD patients positive for ADAs and 239 IBD patients negative for ADAs to anti-TNFα.

For the replication of HLA-DRB1*03 only the UMCG cohort was assessed, with 56 IBD patients positive for ADAs and 143 IBD patients negative for ADAs to anti-TNFα.

Immunochip imputation and quality control

Immunochip genome-wide imputation was done using the Michigan imputation server.[12] Data were first pre-phased with SHAPEIT2,[13] then imputed using IMPUTE2[14] with Haplotype Reference Consortium (r1.1) as the reference panel. An imputation info score > 0.80 was applied.

Stringent quality control was performed. Using default settings in VCFtools, all variants with a call rate of ≤ 0.99 or MAF ≤ 0.001 and variants with a Hardy-Weinberg equilibrium P value ≤ 0.0001 were removed. Duplicated and related samples were also removed.

After these filtering steps, the final UMCG cohort contained 56 samples positive for ADAs and 143 samples negative for ADAs and genotype information for 1,332,743 variants. The final Leuven cohort contained 65 samples positive for ADAs and 96 samples negative for ADAs and genotype information for 1,189,042 variants.

Genome-wide association meta-analysis

A genome-wide association meta-analysis using inverse-variance weighting was carried out, combining the two cohorts. Association testing was performed using PLINK. Genetic variants were selected if they had a P value < 10^{-5} in the meta-analysis, an association signal in the same direction, and a statistically significant P value of < 0.05 in both cohorts. Lead single nucleotide polymorphisms (SNPs) were determined by pairwise linkage disequilibrium using SNP Annotation and Proxy Search (SNAP).[15] A P value < 10^{-8} was considered genome-wide significant and a P value < 10^{-5} was considered suggestive.

Functional annotation of SNPs and genes

Locuszoom was used to construct regional association plots.[16] Lead SNPs (P value < 10^{-5}) identified with SNAP were selected for further characterization by looking into their association with complex traits and effect on gene expression (eQTL). The Genome Aggregation Database (gnomAD) was used for the functional annotation of SNPs,[17] as were the results from the Genotype-Tissue Expression (GTEx) Consortium, for which multiple tissue eQTL are available.[18] In addition, we used an in-house-developed tool called GeneNetwork.[19] GeneNetwork uses results from the Gene Expression Omnibus microarray to predict pathways and function against other biological databases such as the Gene Ontology and Reactome. Ensembl was used to annotate genes and to predict regulatory functions.[20]

Results

HLA regions associated with ADAs to anti-TNFα

The genome-wide association meta-analysis was carried out by combining the two cohorts, with a total of 121 IBD patients positive for ADAs and 239 IBD patients negative for anti-TNFα ADAs. This meta-analysis showed a suggestive association signal for variant 6:32608610 located in the *HLA-DQA1* locus ($P = 2.3 \times 10^{-5}$, Odds ratio (OR) = 2.09). Furthermore, there was a suggestive association signal for variant 6:32566021 ($P = 5.5 \times 10^{-5}$, OR = 1.98), which is located in an mRNA, AF522251, that is involved in the expression of *HLA-DQB1*/*DRB1* (Table 1A).

For the replication of HLA-DQA1*05, both cohorts (UMCG and Leuven) were combined for a total of 121 IBD patients positive for anti-TNFα ADAs and 239 IBD patients negative for these ADAs. The HLA-DQA1*05 allele ($P = 1.0 \times 10^{-3}$) was associated with the development of ADAs, thus replicating the HLA-DQA1*05 association found by Sazonovs *et al* (Table 1B).

For the replication of HLA-DRB1*03, only the UMCG cohort was assessed, with 56 IBD patients positive for anti-TNFα ADAs and 143 IBD patients negative for these ADAs. The HLA-DRB1*03 allele was not associated with ADAs in IBD patients (UMCG patients) treated with infliximab or adalimumab (P = 0.40), thus we were not able to replicate the HLA-DRB1*03 association found by Billiet *et al* (Table 1C).

Non-HLA regions associated with ADAs to anti-TNFα

A genome-wide association meta-analysis was performed for 121 IBD patients positive for anti-TNFα ADAs and 239 IBD patients negative for these ADAs. We did not identify any signals at genome-wide significance level (P value $< 1.0 \times 10^{-8}$), but did identify eight lead SNPs with a suggestive P values $< 10 \times 10^{-5}$, and these are listed along with their candidate genes in Table 2.

Genetic variant rs1984590 ($P = 3.6 \times 10^{-5}$, OR = 0.43) is located in the *FAM107B* gene (Figure 1), which is involved in Toll Like Receptor (TLR) cascades ($P = 3.9 \times 10^{-4}$) and in the TGF-β receptor signalling pathway ($P = 1.3 \times 10^{-3}$).[19] Another interesting non-HLA signal is genetic variant rs11167828 ($P = 7.6 \times 10^{-5}$, OR = 2.18), which is located near the minor histocompatibility protein HB-1 (*HMHB1*) gene (Figure 2). This protein is a precursor of the minor histocompatibility antigen HB-1 (mHag HB-1), an immunogenic peptide that is presented on the cell surface by MHC class I HLA-B44 and that can generate an immune response after recognition by specific T cells. A Locus plot of the other lead SNPs and their candidate genes can be found in **Supplementary File 1**.

Table 1 HLA regions associated with anti-drug antibodies to anti-TNFα.
A: Results from the genome-wide association meta-analysis by combining UMCG and Leuven cohort, with a total of 121 IBD patients positive for the development of antibodies and 239 IBD patients negative for anti-TNFα antibodies.

Chr	BP Position	SNP	A1	Meta-analysis P value	OR	Q	I	UMCG P value	OR	Leuven P value	OR	INFO score	Candidate Genes
6	32608610	6:32608610	C	2.31×10^{-5}	2.0939	0.473	0	0.01081	1.853	0.0004482	2.381	0.91	*HLA-DQA1*
6	32566021	6:32566021	G	5.53×10^{-5}	1.9422	0.9092	0	0.004688	1.907	0.003573	1.98	0.93	mRNA AF522251

Candidate genes are identified by one of the gene prioritization methods (GnomAd, GTEx, GeneNetwork and Ensemble).
BP: Basepair; Chr: chromosome; OR: odds ratio; SNP: single nucleotide polymorphisms; A1: the effect (OR) with respect to the A1 allele; I: Heterogeneity I^2 percentage; Q: P value for Cochrane's Q statistic, INFO score: imputation score, cut off > 0.80.

B: Results for the association analyses for the HLA-DQA1*05 allele by combining UMCG and Leuven cohort, with a total of 121 IBD patients positive for the development of antibodies and 239 IBD patients negative for anti-TNFα antibodies.

	Number of allele carriers (ADA +) (%)	Number of allele carriers (ADA -) (%)	P value	Coefficient estimate
HLA-DQA1*05	64 (53%)	82 (34%)	1.0×10^{-3}	0.15

Heterozygous allele carriers occur twice, and homozygous allele carriers occur only once.
ADA: anti-drug antibody; IBD: inflammatory bowel disease.

C: Results for the association analyses for the HLA-DRB1*03 allele in het UMCG cohort, with 56 IBD patients positive for the development of antibodies and 143 IBD patients negative for anti-TNFα antibodies.

	Number of allele carriers (ADA +) (%)	Number of allele carriers (ADA -) (%)	P value	Coefficient estimate
HLA-DRB1*03	15 (21%)	30 (27%)	0.40	0.08

Heterozygous allele carriers occur twice, and homozygous allele carriers occur only once.
ADA: anti-drug antibody; IBD: inflammatory bowel disease.

Table 2 Results from the genome-wide non-HLA association meta-analysis by combining UMCG and Leuven cohort, with a total of 121 IBD patients positive for the development of antibodies and 239 IBD patients negative for anti-TNFα antibodies.

Chr.	BP Position	SNP	A1	Meta-analysis P value	OR	Q	I	UMCG P value	OR	Leuven P value	OR	INFO score	Candidate Genes
11	114847733	rs10736477	G	1.16×10^{-5}	0.4627	0.9525	0	0.001219	0.458	0.002431	0.4677	0.82	-
3	37620481	rs6419833	A	1.69×10^{-5}	2.0545	0.6335	0	0.004063	1.91	0.001009	2.242	0.a83	*ITGA9*
10	14750252	rs1984590	A	3.65×10^{-5}	0.4318	0.2372	28.44	0.02199	0.5367	0.0001761	0.331	0.91	*FAM107B*
2	120043214	rs3795891	G	4.57×10^{-5}	0.5012	0.4958	0	0.000833	0.4437	0.01238	0.5591	0.98	*C2orf76*
3	179795773	rs6771471	T	5.07×10^{-5}	2.0417	0.7538	0	0.007005	1.934	0.001983	2.16	0.81	-
5	140780321	rs6880273	A	5.71×10^{-5}	2.0527	0.9403	0	0.004337	2.026	0.003784	2.081	0.94	*PCDHG* and subfamily
11	114785756	rs2513473	A	6.14×10^{-5}	0.5008	0.9365	0	0.002899	0.4943	0.006283	0.5081	0.88	-
5	143160225	rs11167828	C	7.65×10^{-5}	2.1757	0.4381	0	0.01426	1.898	0.001029	2.579	0.90	*HMHB1*

Candidate genes are identified by one of the gene prioritization methods (GnomAd, GTEx, GeneNetwork and Ensemble).
BP: Basepair; Chr: chromosome; OR: odds ratio; SNP: single nucleotide polymorphisms; A1: the effect (OR) with respect to the A1 allele; I: Heterogeneity I^2 percentage; Q: P value for Cochrane's Q statistic, INFO score: imputation score, cut off > 0.80.

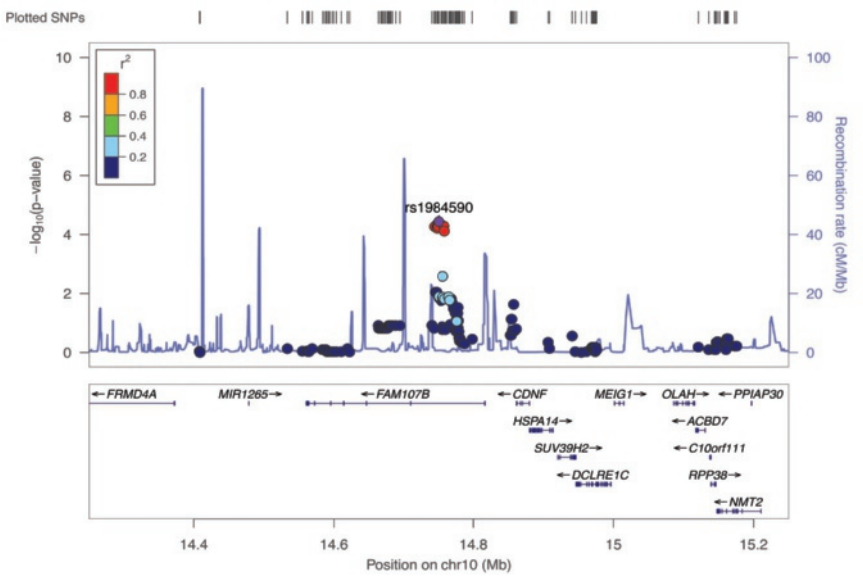


Figure 1 Regional association plot of genetic variant rs1984590 showing a suggestive genetic association for ADA formation (purple triangle). Filled in circles are genotyped or imputed SNP's from the Immunochip. The colour illustrates linkage disequilibrium with the associated SNP

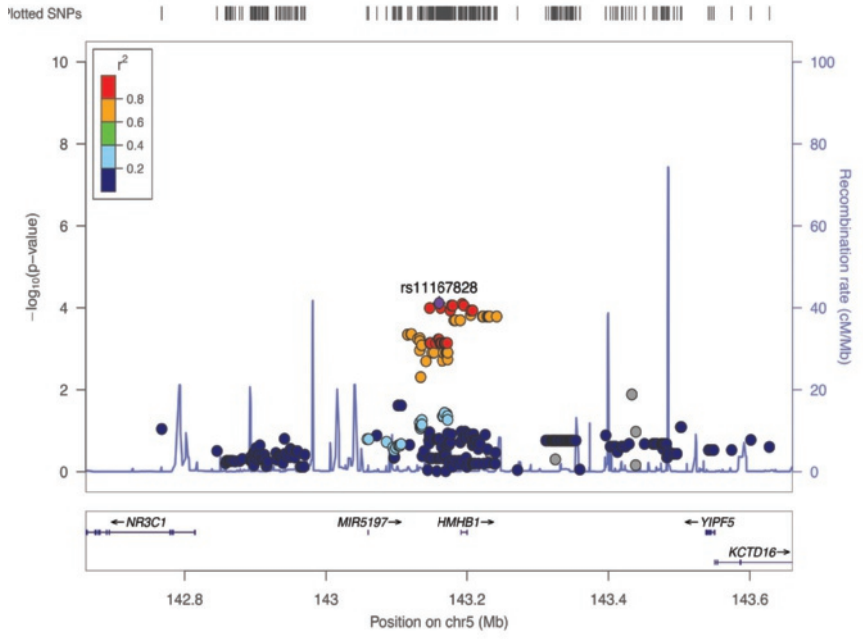


Figure 2 Regional association plot of genetic variant rs11167828 showing a suggestive genetic association for ADA formation (purple triangle). Filled in circles are genotyped or imputed SNP's from the Immunochip. The colour illustrates linkage disequilibrium with the associated SNP

Discussion

In this study we aimed to replicate known HLA regions (HLA-DQA1*05 and HLA-DRB1*03) associated with the development of ADAs to anti-TNFα in IBD patients, as well as to identify novel (non-HLA) genetic regions associated with the development of these ADAs. We were able to replicate the association of the HLA-DQA1*05 allele with ADA formation to anti-TNFα in 121 IBD patients positive for ADAs and 239 IBD patients negative for anti-TNFα ADAs, but were not able to replicate the association with the HLA-DRB1*03 allele described by Billiet *et al.*[5] Our non-HLA genome-wide association meta-analysis in 121 IBD patients with anti-TNFα ADAs and 239 IBD patients negative for these ADAs did not show any signals at genome-wide significance level, but did show eight suggestive association signals in non-HLA regions.

We were able to replicate the association of the HLA-DQA1*05 allele with ADA formation in our cohort, but not the association of the HLA-DRB1*03 allele. Although the HLA-DRB1*03 allele has been associated with ADA formation, it also has been shown to be correlated with the HLA-DQA1*05 allele. We were not able to distinguish whether the HLA-DQA1*05 allele might have driven the signal of the HLA-DRB1*03 allele association, has an independent effect or that they drive the same phenotypic association. Data on genetic associations predicting the development of ADAs to anti-TNFα in IBD patients are scarce. However, a study in multiple sclerosis identified an protective association of the HLA-DRB1*03:01-DQA1*05:01-DQB1*02:01 haplotype associated with the development of ADAs in patients treated with anti-interferon-beta (anti-IFNβ).[21]

This is the first study to assess non-HLA genetic regions associated with the development of ADAs to anti-TNFα in IBD patients. Our genome-wide association meta-analysis showed eight suggestive association signals in non-HLA regions. Genetic variant rs1984590 is located in the *FAM107B* gene, which is involved in the pathway of Toll Like Receptor (TLR) cascades and TGF-β receptor signalling.[19] Active TGF-β binds to its receptor and regulates mucosal immune reactions through the TGF-β signalling pathway, playing a role in the pathogenesis of IBD.[22] Another interesting non-HLA signal is genetic variant rs11167828, which is located near the minor histocompatibility protein HB-1 (*HMHB1*) gene. Minor histocompatibility antigens play a role in the adaptive immune system, as has been shown in graft-versus-host disease in HLA-matched allogeneic hematopoietic cell transplantation.[23,24]

Literature on the genetic background of the development of ADAs in IBD patients treated with anti-TNFα is scarce and conflicting. One of the main reasons for this are the small sample sizes caused by a lack of genotype data, missing information about antibody status, and the fact that only a small subgroup of patients develop ADAs. These small sample sizes make it more difficult to identify signals at genome-wide significance level. As the current cohort was underpowered, we also could not identify any signals at a genome-wide significance level. In our study, HLA-haplotypes

were derived from SNP-based imputation, but more information could have been obtained by using next generation sequence technology or by HLA typing through polymerase chain reaction with sequence-specific primers (PCR-SSP). Differences in allele frequencies between populations in the HLA region could contribute to the discrepancy in genetic association signals in the HLA between different populations. Unfortunately, information on differences in allele frequencies between the Belgian and Dutch populations regarding HLA-DRB1*03 and HLA-DQA1*05 are not available through http://www.allelefrequencies.net. Furthermore, this study also used two different assays and two different time points for measuring ADAs, and this could also have influenced our results.

Understanding the underlying factors that contribute to the development of ADAs to anti-TNFα is important because it would allow us to pre-select those patients who will benefit from treatment. Studies on genetic association with the development of ADAs are very relevant, as their findings could be applied to the development of immunogenicity assays based on patient HLA haplotype. These assays could be used in clinical practice as predictive biomarkers for the development of ADAs, thus allowing personalized treatment for individual patients.

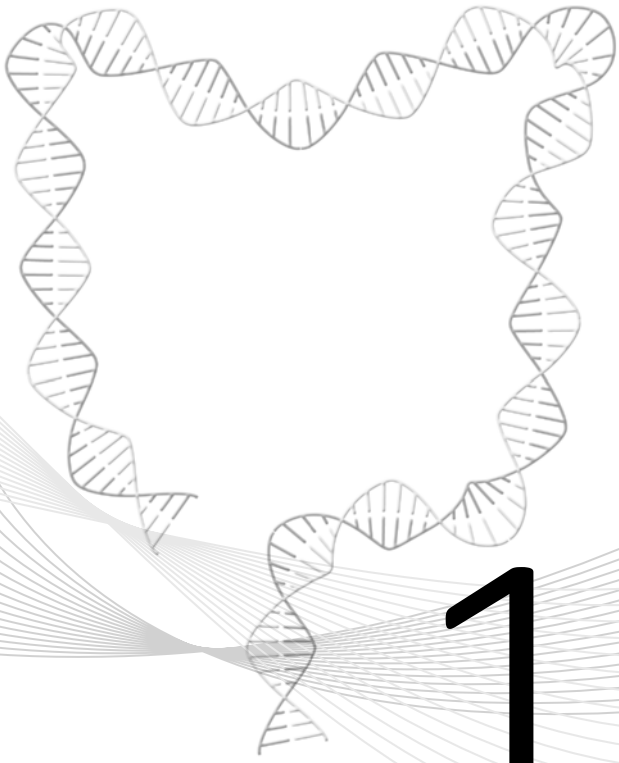
In conclusion, we were able to replicate the HLA-DQA1*05 allele associated with ADA formation to anti-TNFα in IBD patients. We also identified eight suggestive association signals in non-HLA regions that need to be replicated in larger cohorts. As the field of IBD is moving towards personalised medicine, it is important to conduct studies that focus on immunogenicity because future immunogenicity assays might serve as biomarkers to predict treatment response to anti-TNFα therapy.

References

1. Pouillon L, Bossuyt P, Peyrin-Biroulet L. Considerations, challenges and future of anti-TNF therapy in treating inflammatory bowel disease. *Expert Opin Biol Ther* 2016;16:1277-90.
2. Danese S, Fiorino G, Reinisch W. Review article: causative factors and the clinical management of patients with Crohn's disease who lose response to anti-TNFα therapy. *Aliment Pharmacol Ther* 2011;34:1-10.
3. Hindryckx P, Novak G, Vande Casteele N, *et al*. Incidence, Prevention and Management of Anti-Drug Antibodies Against Therapeutic Antibodies in Inflammatory Bowel Disease: A Practical Overview. *Drugs* 2017;77:363-77.
4. Sazonovs A, Kennedy NA, Bewshea C, *et al*.; PANTS Investigator Consortium AT. OP013 HLA-DQA1 contributes to the development of antibodies to anti-TNF therapy in Crohn's disease. 2018. https://www.ecco-ibd.eu/publications/congress-abstract-s/abstracts-2018/item/op013-hla-dqa1-contributes-to-the-development-of-antibodies-to-anti-tnf-therapy-in-crohn-x2019-s-disease-2.html
5. Billiet T, Vande Casteele N, Van Stappen T, *et al*. Immunogenicity to infliximab is associated with HLA-DRB1. *Gut* 2015;64:1344-5.
6. Jostins L, Ripke S, Weersma RK, *et al*. Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* 2012;491:119-24.
7. Shah TS, Liu JZ, Floyd JAB, *et al*. optiCall: a robust genotype-calling algorithm for rare, low-frequency and common variants. *Bioinformatics* 2012;28:1598-603.
8. Purcell S, Neale B, Todd-Brown K, *et al*. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet* 2007;81:559-75.
9. Deelen P, Bonder MJ, van der Velde KJ, *et al*. Genotype harmonizer: automatic strand alignment and format conversion for genotype data integration. *BMC Res Notes* 2014;7:901.
10. Danecek P, Auton A, Abecasis G, *et al*. The variant call format and VCFtools. *Bioinformatics* 2011;27:2156-8.
11. Zheng X, Shen J, Cox C, *et al*. HIBAG-HLA genotype imputation with attribute bagging. *Pharmacogenomics J* 2014;14:192-200.
12. McCarthy S, Das S, Kretzschmar W, *et al*. A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet* 2016;48:1279-83.
13. Delaneau O, Zagury J-F, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods* 2012;10:5-6.
14. Howie B, Marchini J, Stephens M. Genotype Imputation with Thousands of Genomes. Chakravarti A, ed. *G3: Genes|Genomes|Genetics* 2011;1(6):457-470.
15. Johnson AD, Handsaker RE, Pulit SL, *et al*. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* 2008;24:2938-9.
16. Pruim RJ, Welch RP, Sanna S, *et al*. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* 2010;26:2336-7.
17. Lek M, Karczewski KJ, Minikel E V, *et al*. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 2016;536:285-91.
18. GTEx Consortium J, Thomas J, Salvatore M, *et al*. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 2013;45:580-5.
19. Gene Network. http://www.genenetwork.nl/
20. Zerbino DR, Achuthan P, Akanni W, *et al*. Ensembl 2018. *Nucleic Acids Res* 2018;46:D754-61.
21. Núñez C, Cénit MC, Alvarez-Lafuente R, *et al*. HLA alleles as biomarkers of high-titre neutralising antibodies to interferon-β therapy in multiple sclerosis. *J Med Genet* 2014;51:395-400.
22. Ihara S, Hirata Y, Koike K. TGF-β in inflammatory bowel disease: a key regulator of immune cells, epithelium, and the intestinal microbiota. *J Gastroenterol* 2017;52:777-87.

23. Larsen ME, Kornblit B, Larsen MV, *et al*. Degree of Predicted Minor Histocompatibility Antigen Mismatch Correlates with Poorer Clinical Outcomes in Nonmyeloablative Allogeneic Hematopoietic Cell Transplantation. *Biol Blood Marrow Transplant* 2010;16:1370-81.

24. Heinold A, Opelz G, Scherer S, *et al*. Role of Minor Histocompatibility Antigens in Renal Transplantation. *Am J Transplant* 2007;8:95-102.



11

Conclusions, Discussion and Future Perspectives

Conclusions

The aim of my thesis was to identify clinical and genetic predictors for specific IBD disease phenotypes.

In order to identify factors associated with a specific disease behaviour, it is necessary that patient phenotypes are described consistently. The Montreal classification is a classification system for sub-phenotypes of both CD and UC that is widely used in the clinic, but for which almost no data on its reliability and reproducibility was available. In **chapter 2**, we therefore validated the Montreal classification among 30 observers with different professions (gastroenterology specialist in IBD, gastroenterologist in training, and IBD-nurses) in 20 de-identified medical records. We found good to excellent inter-observer agreement for all Montreal items except disease severity in UC, for which agreement was poor. This validation step was important for **chapter 3**, which provides an overview of the phenotypic characteristics currently present in the IBD Parelsnoer cohort, where Montreal classification is an important phenotype. This cohort was founded by a collaboration of the eight University Medical Centers (UMCs) in the Netherlands, and consists of 3388 IBD patients. Chapter 3 describes the design and baseline characteristics for 225 IBD-related items present in the IBD Parelsnoer cohort. This chapter also describes the IBD Parelsnoer cohort's potential to facilitate clinical and basic science and to improve clinical care. Data presented in chapter 3 is used in subsequent chapters to assess clinical parameters associated with phenotypic differences in patients with IBD.

As patients with IBD can suffer significantly from disease symptoms or flares, they are at risk for work disability. In **chapter 4**, I assessed potential (clinical) risk factors for work disability in IBD, and found that work disability was associated with female sex, a lower education level, extraintestinal manifestations, an age > 55 years, an age > 40 years at diagnosis, a disease duration > 15 years, smoking, surgical interventions, anti-TNFα use and immunomodulator use. In CD patients, long-term full work disability (defined as > 80% work disability for > 2 years) was associated with a lower education level. In UC patients, long-term full work disability was associated with complications (osteopenia, thromboembolic events).

The incidence of IBD in developing countries may be increasing due to industrialization and Western lifestyle changes, but these changes cannot explain the phenotypic heterogeneity within and between regions. In **chapter 5**, I therefore assessed the impact of ethnicity and country of birth on IBD phenotype. What I found was that patients of non-Caucasian descent (non-CEU) with CD more often had upper gastro-intestinal disease (L4) and anal stenosis than patients with CD from West- and Central-European Caucasian descent (CEU). Non-CEU patients with IBD used more anti-TNFα agents and immunomodulators than CEU patients. The most interesting finding

was related to country of birth: non-CEU IBD patients born in Europe were diagnosed at a younger age than non-CEU IBD patients born outside Europe, implying that Western lifestyle might trigger IBD onset more early in life.

Treatment strategies in IBD are still applied equally to male and female patients. If we want to move further towards personalized treatment it is important to be aware of differences in IBD disease course and disease phenotypes between the sexes. In **chapter 6**, I focused on differences between males and females with IBD, comparing phenotype, clinical manifestations, disease course, medical treatment and other healthcare consumption. What we found was that early onset CD (< 16 years) was more frequently seen in males, male patients with CD more often had ileal disease and male patients more often underwent small bowel and ileocaecal resection. Male patients with CD also more frequently suffered from osteopenia and used prednisone more often. Extraintestinal manifestations (EIMs), on the other hand, were more often observed in female IBD patients. IBD-specific healthcare costs did not differ between male and female IBD patients.

In Part II of this thesis I identified genetic risk loci associated with disease behaviour. Part II starts with a review (**chapter 7**) describing the clinical presentation of IBD and gives an overview of the progress that has been made in the field of IBD genetics, from linkage and candidate studies to GWAS and Immunochip studies. Chapter 7 then describes the genetic and biological pathways of IBD and its overlap with other immune-mediated disease. In the last part of the review, I focus on genetic findings and how we can translate these findings to clinical practice.

Hidradenitis suppurativa (HS) is a chronic inflammation of the apocrine glands that is often followed by sinus tract formation and scarring. Although HS is not yet a well-recognized EIM in IBD, the prevalence of HS in IBD is much higher than in the general population, which suggests a shared pathogenesis. In **chapter 8**, I identify genetic and clinical parameters associated with HS in IBD. Female sex, patients with CD, smoking, a higher body mass index, and younger age were independently associated parameters for HS in IBD. The within-cases allelic association analysis revealed a suggestive protective association with the *ELOVL7* gene and a suggestive risk association with the genes *SULT1B1* and *SULT1E1* for HS in the context of IBD.

The need for surgery for fibrostenotic disease in patients with CD is an indicator of a severe disease course. In a within-cases allelic association analysis, we identified the *WWOX* gene as a disease-modifying gene associated with recurrent fibrostenotic CD disease (**chapter 9**). Functional studies showed an enhanced colonic expression of Transforming Growth Factor-beta (TGF-β) in *WWOX* risk-allele carriers, supporting our hypothesis of *WWOX* as having a role in fibrosis formation.

Anti-TNFα agents are widely used for inducing and maintaining clinical remission in patients with CD. However, some patients develop anti-drug antibodies (ADAs) to anti-TNFα, resulting in loss of response. In **chapter 10,** I identified genetic risk variants that play a role in the development

of ADAs to anti-TNFα (infiximab and adalimumab) by comparing IBD patients who developed anti-TNFα ADAs to IBD patients without these ADAs. I was able to replicate the association of the HLA-DQA1*05 allele associated with ADA formation to anti-TNFα. However, I was not able to replicate the association with the HLA-DQB1*03 allele, which is also known to be associated with anti-TNFα ADAs. I further identified eight suggestive association signals in non-HLA regions, which will need to be replicated in larger cohorts.

Discussion and Future Perspectives

Future of patient care - Precision medicine

The goal of precision medicine is to tailor medical decisions or practice to the needs of an individual patient using epidemiological, genetic or other molecular or cellular analysis. Precision medicine should create the ability to classify individuals into subpopulations, to weigh the chance of an individual developing a particular disease, to predict that individual's disease course/prognosis, and to know which medication or intervention that patient will benefit most from. In precision medicine the right group of patients will benefit from the right treatment at the right moment, while other patients will be protected from the side-effects or from adverse outcomes.

The concept of precision medicine is not novel. General practitioners have been using precision medicine in their daily practice for some time now. For example, the ten-year-risk of a cardiovascular event is assessed for each individual patient depending on the patient's clinical characteristics and parameters such as age, sex, smoking status, systolic blood pressure, and TC/HDL ratio. The combination of these risk factors predicts the risk of a cardiovascular event for the individual patient in the coming ten years. Depending on this risk score, a patient will then receive either solely lifestyle advice or medical treatment with statins. This is an example of a precision medicine model in which only clinical parameters have been included. Inclusion of other layers, for example genetic status, could further improve the performance of risk prediction models.

In 2015 Barack Obama, former president of the United States (US) announced the launch of Precision Medicine Initiative later called "All of Us" Research Program. The National Institute of Health (NIH) leads this initiative to build a national large-scale research enterprise with extensive information about lifestyle, environment and genetics. Participants will be from diverse ancestral, racial/ethnic, geographic, social and economic backgrounds, and from all age groups and health statuses, reflecting the diversity of the US population. The idea is to enrol more than one million US citizens. The main objective of the "All of Us" Research Program is to enable scientific research for common and rare disease and increase our knowledge of an individual's chances of staying healthy throughout life. Although we don't have a national Biobank initiative of this size or scope

in the Netherlands, we do have Lifelines, which is comparable to the "All of Us" Research Program in several ways. Lifelines is a population-based prospective cohort that includes more than 167,000 participants from the three Northern provinces of the Netherlands. This group of participants consists of three generations and will be followed-up over the coming thirty years. The main objective of Lifelines is to enable research to monitor the process of ageing and to identify factors related to health and disease. Detailed phenotypic and environmental factors are collected by questionnaire every 18 months. Health status parameters are measured every 5 years. For 10% of the participants, genetic information is available. In addition to this large population-based cohort, 1500 participants were asked to participate in the Lifelines DEEP cohort. Where Lifelines is comparable to "All of Us" Research Program, Lifelines DEEP takes it to the next level by including many more molecular layers. For the 1500 Lifelines DEEP participants, exhaled air (analysis of volatile organic compounds), faecal samples (microbiome and biomarker assessment), and additional blood (genetics, methylation and transcriptomics analyses) were collected. This rich dataset enables us as researchers to combine several molecular layers and address questions related to ageing from different perspectives (demographics, genotype, microbiome, transcriptome and methylation).

Another example of a cohort suitable for translational research is the IBD Parelsnoer cohort that I worked with during my PhD project. The IBD Parelsnoer cohort aims to facilitate the discovery of predictors (both epidemiological risk factors and biomarkers) for individual disease course and treatment response in IBD. Major strengths of the IBD Parelsnoer cohort are its prospective design, its extensive uniform information model comprising 225 data items, and the participation of all eight UMCs in the Netherlands. In addition, biomaterials such as serum, DNA, stool samples and, if available, biopsies from endoscopy and resection tissue are collected. This enables researchers to discover new biomarkers by the integration of molecular and clinical phenotypes. Another strength of the IBD Parelsnoer cohort is its wide collaboration with international IBD research groups.

An example of how precision medicine has made great progress by adding molecular levels comes from the field of oncology. Almost all tumours are now screened on a molecular level, producing information that allows medication to be adjusted to the individual patient and allows for chemotherapy—which is both physically and financially taxing—only to be prescribed to patients who will benefit from it. Severe side-effects in patients who don't benefit from these drugs and high medical costs can thus be prevented.

Precision medicine in IBD and other complex diseases

For complex multifactorial diseases such as IBD, precision medicine is still in its infancy, but it is starting to evolve. Precision medicine in IBD begins with the identification of clinical parameters that could influence IBD risk. Having had an appendectomy, for example, is protective against the development of UC, while smoking can increase the risk of developing CD. Since the influence

of genetic factors became evident in IBD (heritability in monozygotic twins), we have added the first molecular layer—genetics—by performing linkage studies and candidate gene association studies. In 2000, we started with GWAS analysis and by 2012 had identified 163 loci associated with IBD risk. However, these genetic risk variants only explained 20% of the genetic disease risk. To identify new genetic risk variants (which probably have much smaller effects) by GWAS, an immense increase in statistical power was needed. As most of the samples available in the Western World had already been included in previous studies, a thousand non-Caucasians samples were added. With the inclusion of these samples, we were able to identify more genetic variants associated with disease risk and bring the total number of genetic loci for IBD risk up to 242. However, increasing the number of risk loci did not add to the explained heritability, which remained at slightly over 20%. Furthermore, these identified genetic risk variants explain only about 10% of the total disease risk, suggesting that environmental and other unidentified molecular factors play a role in the pathogenesis of IBD. Although finding new genetic risk variants associated with IBD was the main goal of GWAS and Immunochip studies, this goal has now shifted towards identifying environmental factors that influence disease risk, and finding genetic variants associated with a specific disease phenotype.

As the name 'precision medicine' indicates, the goal is to create an algorithm that can predict disease course, the patient's individual response to therapy and any adverse drug response, as opposed to solely determining risk to develop IBD. Therefore, the second part of my thesis focuses on the identification of genetic variants associated with specific disease phenotypes (hidradenitis suppurativa in IBD in Chapter 8, and recurrent fibrostenotic CD in Chapter 9) and immunogenicity to anti-TNFα (antibody development against anti-TNFα treatment in Chapter 10). Unfortunately, my sample size in Chapter 8 and Chapter 10 was too small to gain enough power to detect an association signal at genome-wide statistical significance level (p-value < 5×10^{-8}). One way to increase power is by adding more samples. Luckily, large consortia are now collaborating to unravel the genetic background of disease behaviour and response to drugs. These international consortia have, for example, found that the *HLA-DQA1-HLA-DRB1* haplotype confers susceptibility to thiopurine-induced pancreatitis. Patients who are homozygotic for genetic variant rs2647087 have a 17% risk of developing pancreatitis after administration of a thiopurine in comparison to heterozygote patients, who have a 9% risk.

In addition to better understanding drug toxicity, we also aim to identify factors that predict drug response. Two new biological drugs, vedolizumab and ustekinumab, have been registered in the Netherlands as drug therapy for refractory disease in IBD. Both target a different pathway than well-known anti-TNFα agents (infliximab, adalimumab, golimumab). Vedolizumab has been shown to be more effective in biological-naïve patients compared to biological-exposed patients. We are now taking mucosal biopsies from biological-naïve patients and anti-TNFα-exposed patients before

and after treatment with vedolizumab. When we have identified which immune cells in the mucosal layer are associated with response to vedolizumab treatment, we can compare these to profiles of mucosal infiltrates present in pre-treatment biopsies. By doing so, we will be able to personalize vedolizumab treatment by selecting those patients who are likely to respond to vedolizumab treatment. An important question remains, however: Will it be cost-effective to determine these risk variants in every patient before starting treatment? What could make this financially feasible is if many specialties in the UMCG together develop a "response to drug" panel that includes variants associated with drug toxicity and variants associated with drug response. Testing every patient with a chronic disease before start of treatment might then be cost-effective.

Thus far, GWAS and Immunochip studies have identified 242 loci associated with IBD risk. Some of these risk variants are disease-specific, but most are shared between IBD and other immune mediated diseases. These associated-variants are unlikely to be the causal because they are quite common in the general population (minor allele frequency > 1%). Moreover, these variants mark genomic loci, and these loci can contain several genes in the human genome that influence disease risk. The effect size of genetic risk variants on disease risk tends to be much smaller in complex disease such as IBD, as compared to Mendelian or oligogenic diseases, where rare genetic variants often confer larger risk effects. An example of an oligogenic disease is very-early onset IBD (VEO-IBD), which is characterized by a severe disease course that is unresponsive to conventional therapy. Whole exome sequencing (WES) has identified mutations in the *IL10RA/B* genes in patients with VEO-IBD. This suggests that these rare genetic variants are highly penetrant and have a large effect size that causes this Mendelian-like or oligogenic form of IBD. VEO-IBD patients with these mutations are generally refractory to standard immunosuppressive therapy, which means that hematopoietic stem cell transplantation should be considered as a potential curative treatment option early in the disease course. Through WES we can identify rare genetic variants with large effect sizes that are only present in a small number of patients, and therefore not detected with genotyping arrays such as the Immunochip. In addition, these severe disease-causing variants are much more likely than low-impact variants to be located in the protein coding part of genes that is covered by WES. WES can help us to identify rare or novel genetic variants in Mendelian-like or oligogenic forms of IBD.

As discussed above, the genetic risk variants identified so far only explain 10% of the risk for disease, which suggests that environmental factors and other unidentified factors also have a role. We now know that the microbiome plays an important role in IBD because there is a clear difference in the microbiome between patients with IBD and healthy controls. This would suggest that it is possible to create a model that can distinguish healthy people from individuals suspected of having IBD by characterizing their gut microbiome. Furthermore, differences in microbiome profiling between IBD patients with active disease and IBD patients in remission could be used to

monitor therapeutic targets. The microbiome is also a relatively easy accessible treatment target. Medication, dietary changes or even faecal microbiome transplantation (FMT) that target the right bacteria could shift the microbiome towards a more 'healthy' gut. FMT has already changed the standard of care for *Clostridium difficile* infection. In IBD, however, the role of FMT is still emerging, and there are no definite conclusions as of yet. For now, the heterogeneity of the disease makes results from different studies difficult to compare and to interpret. For the future, with the right selection of microbial consortiums, microbiome profiling could be a promising development for 'precision medicine'. Other upcoming research fields, such as transcriptomics and methylation studies, are beyond the scope of this thesis. Hopefully, in the coming years, it will be possible to take these layers of molecular information into account in the precision medicine model.

To conclude, IBD is a complex disease. Environmental factors, the microbiome and genetic factors all have an effect on disease risk, disease behaviour, and response to therapy. Combining these factors into risk models for the individual patient is the goal of precision medicine, a goal that will hopefully become feasible in the coming decade.

Precision medicine in this thesis

Throughout this thesis I have used data from the IBD Parelsnoer cohort. Its prospective design, extensive phenotype collection and standardized biomaterial collection, make IBD Parelsnoer very suitable for translational research. With the data available in this biobank, it will be feasible to identify modulators that influence disease course. Biomarker discovery research to assess response to treatment or adverse outcomes is also possible. The data collection process is the main focus of future improvement. Currently, clinical information is collected during each patient visit. Data obtained through web-based questionnaires at standard time points (for example every three months) would be a highly valuable addition. These Patient Reported Outcomes Measurements (PROMS) could give us more insights into time-to-event related questions. For example, if a patient presents with a flare at the patient clinic, we could examine if their PROM shows any signs that pointed towards the development of this flare, e.g. an increase in diarrhoea frequency. This would be an even more reliable way to identify modulators that influence the disease course.

Another important point that needs to be addressed is treatment target goals in IBD. For decades clinic care and clinical trials have focused on the induction of clinical response and maintaining clinical remission. However, treatment strategies focused on clinical remission have not altered the disease course of IBD. Therefore, it is important that we look beyond just the treatment of symptoms toward achieving better outcomes by preventing bowel damage and disability. Mucosal healing and deep remission are currently suggested as good candidates for disease modification, but it has not yet been proven that achieving and maintaining deep remission will alter the disease course of IBD. I think that the IBD Parelsnoer cohort is a very suitable resource

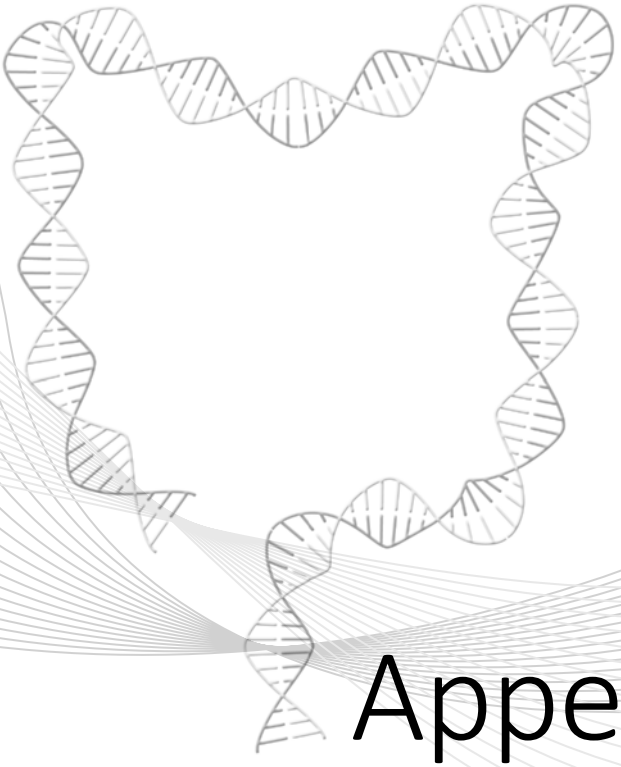
for addressing this 'treatment target' goal, and that we should use this as a clinical endpoint when we assess modulators that can influence the disease course in IBD.

I was the first person to get the opportunity to work with the data available in the IBD Parelsnoer cohort, and this has taught me several important lessons. While working with the data, it became evident that clear definitions are of the utmost importance, as are internal data checks. For example, a clinician should not be able to fill out a Harvey Bradshaw Index for a patient with UC, as this index is only meant for CD patients. To make the data more reliable, we should be able to go back to the patient records. On the other hand, information security guidelines such as the international ISO 27.001 guarantee patient privacy, which is of utmost importance. The IBD Parelsnoer cohort collects phenotypes and biomaterials in a standardized manner, making international collaboration feasible. The IBD Parelsnoer cohort is also part of the Biobanking and Biomolecular Resources Research Infrastructure of the Netherlands (BBMRI-NL), which is the Dutch national node of BBMRI-ERIC, the largest research infrastructure project in Europe. The IBD Parelsnoer cohort is quite young and has the opportunity to grow. I think that choosing to add more molecular layers to the IBD Parelsnoer cohort, e.g. microbiome data, is a valid alternative to increasing the sample size. I believe that adding more molecular layers will help us built a precision medicine logarithm model. I think the IBD Parelsnoer cohort is a great initiative that creates opportunities for research.

Conclusions of this thesis

- The Montreal classification is a reliable tool for the assessment of sub-phenotypes in IBD with the exception of disease severity in UC.
- Female sex, a lower education level, extraintestinal manifestations, an age > 55 years, an age > 40 years at diagnosis, a disease duration > 15 years, smoking, surgical interventions, anti-TNFα use, and immunomodulator-use are associated with work disability in IBD.
- A lower education level in patients with CD and complications (osteopenia, thromboembolic event) in patients with UC are associated with long-term full work disability (> 80% work disability for > 2 years).
- Compared to patients of West- and Central-European Caucasian descent, patients of non-Caucasian descent more often have upper gastro-intestinal disease and anal stenosis and are prescribed more anti-TNFα and immunomodulators.
- Patients of non-Caucasian descent born in Europe are diagnosed at a younger age than patients of non-Caucasian descent born outside Europe.
- Early onset CD (< 16 years) is more frequently seen in males.
- Male patients with CD are more likely to have ileal disease and more often undergo small bowel and ileocaecal resection than females.

- Male patients with CD more frequently suffer from osteopenia and use prednisone more often than female CD patients.
- Extraintestinal manifestations are more often observed in female IBD patients.
- IBD-specific healthcare costs do not differ between male and female IBD patients.
- Female sex, patients with CD, smoking, a higher body mass index, and younger age are independently associated parameters for hidradenitis suppurativa in IBD.
- For hidradenitis suppurativa in IBD, we found a suggestive protective association with *ELOVL7* and suggestive risk association with the genes *SULT1B1* and *SULT1E1*.
- The *WWOX* gene is a disease-modifying gene associated with recurrent fibrostenotic CD disease.
- I have replicated the association of the HLA-DQA1*05 allele previously associated with anti-drug antibody (ADA) formation to anti-TNFα in IBD patients.
- I was not able to replicate the association with the HLA-DRB1*03 allele previously associated with ADA formation to anti-TNFα in IBD patients.
- A genome-wide association meta-analysis identified eight suggestive association signals in non-HLA regions associated with anti-TNFα ADAs in IBD patients.



Appendix

Samenvatting, Dankwoord,
Curriculum vitae and List of publications

Samenvatting

Inflammatoire darmziekten ("inflammatory bowel diseases" of "IBD"), zoals de ziekte van Crohn (ZvC) en colitis ulcerosa (CU), zijn chronische ontstekingsziekten van het maagdarmkanaal. Deze ontsteking wordt waarschijnlijk veroorzaakt door een disbalans tussen de bacteriën aanwezig in het maagdarmkanaal en het immuunsysteem. De prevalentie van IBD is het hoogst in Europa en Noord-Amerika. In Nederland is de prevalentie 830 per 100,000 personen.

IBD behoort tot de complexe ziekten, wat inhoudt dat zowel omgevingsfactoren als genetische factoren van invloed zijn op het ontstaan van IBD. Een aantal omgevingsfactoren zijn al bekend. Verwijdering van de blindedarm middels een operatie werkt bijvoorbeeld beschermend voor het ontstaan van CU. Roken daarentegen vergroot de kans op het ontstaan van de ZvC. Het natuurlijk beloop van IBD kent periodes van (langdurige) remissie, afgewisseld door periodes van opvlamming van de ziekte. Klachten die patiënten met IBD kunnen ervaren zijn onder andere buikpijn, (bloederige) diarree, koorts en afvallen.

De lokalisatie in het maagdarmkanaal, de ernst en het beloop van de ziekte kunnen sterk variëren tussen patiënten met IBD. Vanwege deze variabele presentatie is het tot op heden nog niet mogelijk om te voorspellen welke patiënt een verhoogd risico heeft op het ontwikkelen van een ernstig ziektebeloop. Om het beloop beter te kunnen voorspellen, is het van belang om klinische factoren te identificeren die geassocieerd zijn met een specifiek IBD-fenotype. Klinische factoren die geassocieerd zijn met een ernstig ziektebeloop zijn: een jonge leeftijd op het moment van diagnose, een fistelende ziekte en het roken bij de ZvC. De eerste twee risicofactoren zouden een reden kunnen zijn om bij een ernstig ziektebeloop te besluiten tot een agressief behandelplan, zoals een chirurgische interventie of te starten met zware medicatie zoals een biological. Naast deze bekende klinische risicofactoren zijn er een aantal klinische factoren waarvan het effect op het fenotype en ziektebeloop in IBD nog nauwelijks is onderzocht, zoals de invloed van etniciteit en geslachtsverschillen. Daarnaast is er de afgelopen jaren veel onderzoek gedaan naar genetische factoren. Genoomwijde associatiestudies (GWAS), waarbij grote aantallen patiënten en gezonde individuen worden gegenotypeerd en vergeleken voor honderdduizenden genetische varianten, zijn erg succesvol gebleken. Het aantal geassocieerde risicogebieden (genetische varianten) op het genoom voor het ontstaan van IBD is 242. Deze studies zijn echter gericht op genetische varianten die geassocieerd zijn met het ontstaan van IBD. Het aantal studies gericht op een specifiek IBD-fenotype, het ziektebeloop of de respons op therapie is maar beperkt.

Het doel is om uiteindelijk te streven naar de ontwikkeling van een algoritme, waarin zowel klinische als genetische risicofactoren zijn opgenomen en waarmee voorspeld kan worden welke patiënt een verhoogd risico heeft op het ontwikkelen van een ernstig ziektebeloop. Door het vroegtijdig aanpassen van behandelstrategieën op basis van deze risicostratificatie, zullen

ongunstige uitkomsten voorkomen kunnen worden. Deze zorg op maat, aangepast aan de individuele patiënt, wordt ook wel 'precision medicine' genoemd. Het doel van mijn proefschrift is daarom het identificeren van klinische en genetische risicofactoren die geassocieerd zijn met een specifiek IBD-fenotype.

Het proefschrift zal zich richten op twee onderwerpen binnen de IBD:

I. Het identificeren van klinische factoren die de fenotypische verschillen in IBD-patiënten kunnen verklaren
II. Het identificeren van genetische varianten die geassocieerd zijn met een specifiek IBD-fenotype

Deel I richt zich op het identificeren van klinische factoren die geassocieerd zijn met een specifiek IBD-fenotype. Om deze factoren te kunnen identificeren, is het van groot belang dat fenotypen op consistente wijze worden beschreven. De Montreal classificatie is een classificatiesysteem dat gebruikt wordt in de kliniek voor het beschrijven van fenotypen voor zowel de ZvC als CU. Er is in de literatuur echter maar weinig bekend over de reproduceerbaarheid en betrouwbaarheid van de Montreal classificatie. In **hoofdstuk 2** valideren we de Montreal classificatie onder dertig waarnemers uit drie verschillende beroepsgroepen (gastro-enterologen gespecialiseerd in IBD, gastro-enterologen in opleiding en IBD-verpleegkundigen). Deze waarnemers scoren de Montreal classificatie in twintig medische dossiers. De reproduceerbaarheid was goed tot uitstekend voor alle Montreal items, met uitzondering van het item ziekte-ernst in CU, deze was slecht. Validatie van de Montreal classificatie is met name belangrijk voor **hoofdstuk 3**, aangezien er in dit hoofdstuk een overzicht wordt gegeven van de aanwezige fenotypen in het IBD Parelsnoer cohort, waarvoor onder andere de Montreal classificatie is gebruikt. Het IBD Parelsnoer cohort is opgericht door een samenwerking van de acht Universitaire Medische Centra (UMC's) in Nederland en bestaat uit 3388 IBD-patiënten. Hoofdstuk 3 is een beschrijving van het design en de opzet van het IBD Parelsnoer cohort. Daarnaast geeft het een overzicht van de baseline karakteristieken voor 225 IBD gerelateerde onderwerpen en beschrijft het de mogelijkheden voor welke doeleinden het IBD Parelsnoer cohort gebruikt zou kunnen worden. Denk aan: klinisch onderzoek, basaal onderzoek en het verbeteren van de klinische zorg. De data beschreven in hoofdstuk drie zullen tevens gebruikt worden in de daarop volgende hoofdstukken, waarin de associatie tussen klinische factoren en fenotypische verschillen bij IBD onderzocht zal worden.

IBD kent een onvoorspelbaar ziektebeloop, waarin patiënten vaak klachten ervaren die van invloed zijn op hun dagelijks functioneren. Patiënten met IBD hebben daardoor een verhoogd risico op arbeidsongeschiktheid. In **hoofdstuk 4** kijk ik naar klinische (risico)factoren die mogelijk geassocieerd zijn met arbeidsongeschiktheid bij patiënten met IBD. Het vrouwelijk geslacht, een laag opleidingsniveau, extra-intestinale manifestaties, leeftijd boven de 55 jaar, leeftijd boven de

40 jaar bij diagnose, een ziekteduur van meer dan 15 jaar, roken, chirurgische ingrepen, anti-TNFα gebruik en het gebruik van immunomodulatoren, waren geassocieerd met arbeidsongeschiktheid. Bij patiënten met de ZvC was een laag opleidingsniveau geassocieerd met langdurige volledige arbeidsongeschiktheid en bij patiënten met CU waren dat complicaties (osteopenie en trombo-embolische events) (> 80% arbeidsongeschiktheid voor > 2 jaar).

De incidentie van IBD in ontwikkelingslanden is de afgelopen jaren toegenomen, meest waarschijnlijk is dit het gevolg van industrialisatie en Westerse leefstijlveranderingen. Echter, dit verklaart niet de fenotypische heterogeniteit binnen en tussen de regio's. In **hoofdstuk 5** onderzoek ik daarom de invloed van etniciteit en geboorteland op het IBD-fenotype. Patiënten met de ZvC van niet-Kaukasische afkomst hadden vaker ziekteactiviteit gelokaliseerd boven de dunne darm (proximaal van het ileum (L4)) en vaker een anale stenose in vergelijking met patiënten met de ZvC van een West- en Centraal-Europese Kaukasische afkomst. IBD-patiënten van niet-Kaukasische afkomst gebruikten bovendien vaker anti-TNFα en immunomodulatoren dan IBD-patiënten van West- en Centraal-Europese Kaukasische afkomst. Een interessante bevinding was de associatie tussen geboorteland en leeftijd van diagnose. IBD-patiënten van niet-Kaukasische afkomst geboren in Europa werden op een jongere leeftijd gediagnosticeerd met IBD dan IBD-patiënten van niet-Kaukasische afkomst geboren buiten Europa. Dit zou kunnen betekenen dat bij blootstelling aan een Westerse leefstijl IBD zich eerder manifesteert.

Bij IBD wordt geen onderscheid gemaakt tussen mannen en vrouwen als het gaat om behandelstrategieën. Aangezien we ons steeds meer richten op 'precision medicine', is het belangrijk om op de hoogte te zijn van de verschillen tussen mannen en vrouwen met IBD als het gaat om ziektelokalisatie, ziektegedrag, medicatiegebruik en zorgkosten. In **hoofdstuk 6** richt ik mij op de klinische verschillen tussen mannen en vrouwen met IBD. De diagnose IBD op jonge leeftijd (< 16 jaar) werd vaker gezien bij mannen met IBD. Mannen met IBD hadden vaker ziekteactiviteit in de dunne darm en ondergingen vaker een dunne darm resectie of ileocecaal resectie dan vrouwen met IBD. Osteopenie werd vaker geconstateerd bij mannen met IBD en zij gebruikten vaker prednisolon. Vrouwen met IBD daarentegen hadden vaker extra-intestinale manifestaties. Er was geen verschil tussen mannelijke en vrouwelijke IBD-patiënten als het ging om IBD-specifieke zorgkosten.

In deel II van mijn proefschrift identificeer ik genetische varianten die geassocieerd zijn met een specifiek IBD-fenotype.

Het tweede deel begint met een review (**hoofdstuk 7**). In dit review wordt de klinische presentatie van IBD beschreven en een overzicht gegeven van de vooruitgang die is geboekt op het gebied van genetisch onderzoek. Beginnend bij de kandidaat en linkage studies tot aan GWAS en Immunochip studies. Verder worden de genetische en biological pathways betrokken in IBD en wordt hun overlap met andere immuun gemedieerde ziekten besproken. In het laatste deel van

het review ligt de nadruk op gevonden genetische associaties in IBD en de mogelijke toepassing hiervan in de klinische praktijk.

Hidradenitis suppurativa (HS) is een chronische ontsteking van de apocriene klieren, met vaak als gevolg fibrose- en littekenvorming. Hoewel HS officieel geen extra-intestinale manifestatie van IBD is, is de prevalentie van HS bij IBD-patiënten vele malen hoger dan in de algemene populatie, wat een gedeelde pathogenese suggereert. In **hoofdstuk 8** onderzoek ik daarom genetische varianten en klinische factoren die mogelijk geassocieerd zijn met HS bij IBD. Uiteindelijk bleken, vrouwelijk geslacht, patiënten met de ZvC, roken, een hogere Body Mass Index en jongere leeftijd, onafhankelijke risicofactoren geassocieerd met HS in IBD. Een genetische associatieanalyse liet twee suggestieve associaties zien op het genoom. Een variant daarvan bevindt zich in het ELOVL7-gen (beschermend effect) en de andere tussen het ***SULT1B1***- en ***SULT1E1***-gen (risicovariant).

Een voorbeeld van een ernstig ziektebeloop is de patiënt met de ZvC met stenoserende ziekte, waarbij chirurgisch ingrijpen noodzakelijk is. In een genetische associatieanalyse, waarin we patiënten met de ZvC met fibrostenoserende ziekte hebben vergeleken met patiënten met de ZvC met enkel ontstekingsactiviteit van de darm, hebben we (in **hoofdstuk 9**) het ***WWOX***-gen geïdentificeerd. Het ***WWOX***-gen is een bekend tumorsuppressorgen, waarvan in de literatuur beschreven staat dat het mogelijk betrokken is bij de vorming van fibrose. Daarnaast toonden functionele studies een verhoogde colonexpressie van Transforming Growth Factor-beta (TGF-β) in dragers van het ***WWOX***-risico-allel, hetgeen onze hypothese ondersteunt dat ***WWOX*** een rol speelt bij de vorming van fibrose.

Anti-TNFα is een biological die veel wordt gebruikt voor het induceren en behouden van remissie bij patiënten met IBD. Sommige patiënten ontwikkelen echter anti-TNFα antilichamen (ADA's), wat resulteert in het verlies van respons. **Hoofdstuk 10** laat een genetische associatieanalyse zien, waarin we IBD-patiënten die anti-TNFα ADA's ontwikkelden hebben vergeleken met IBD-patiënten zonder deze ADA's. Een al bekende genetische associatie (HLA-DQA1*05-allel) met anti-TNFα ADA's konden we repliceren, een andere bekende associatie met het HLA-DRB1*03-allel werd echter niet gerepliceerd. Daarnaast liet de genetische associatieanalyse acht suggestieve associaties zien in niet-HLA-regio's, die in grotere cohorten met meer statistische power opnieuw onderzocht zullen moeten worden.

Het doel van dit proefschrift was het identificeren van klinische en genetische risicofactoren die geassocieerd zijn met een specifiek IBD-fenotype. Door het identificeren van deze risicofactoren, maar ook van andere moleculaire of cellulaire factoren kunnen algoritmes gericht op 'precison medicine' verder ontwikkeld worden. In de toekomst zullen deze tools het vermogen hebben om de ziekte en de prognose van het individu te voorspellen. Tevens kan beter voorspeld worden bij welke medicatie of interventie de patiënt het meest baat heeft. In 'precison medicine' zal de juiste

groep patiënten profiteren van de juiste behandeling op het juiste moment, terwijl daarentegen andere patiënten worden beschermd tegen de nadelige gevolgen ervan, zoals de bijwerkingen.

Dankwoord

Tijdens mijn bachelor wilde ik graag onderzoek doen. Via de kinder-MDL kwam ik bij de MDL terecht, waar ik onderzoek ging doen onder supervisie van Prof. dr. R.K Weersma en Prof. dr. G. Dijkstra. Na drie jaar fulltime onderzoek en een jaar onderzoek naast mijn werkzaamheden als arts-assistent, is mijn proefschrift eindelijk klaar. Dit was natuurlijk niet gelukt zonder de hulp van velen, daarom wil ik graag de volgende mensen bedanken.

Allereerst wil ik alle IBD-patiënten bedanken voor hun belangeloze bijdrage; zonder jullie had 'the Dutch IBD Biobank' niet bestaan. Verder wil ik het Parelsnoer Instituut en de Initiative on Crohn and Colitis (ICC) bedanken voor het opzetten van deze unieke IBD-biobank. Daarnaast wil ik iedereen bedanken die geholpen heeft met het monnikenwerk dat gepaard ging met het opbouwen van deze database.

Prof. dr. R.K Weersma, beste Rinse, ik heb mij de afgelopen vier jaren geen fijnere supervisor kunnen wensen. Door jouw efficiëntie en kennis is het mij gelukt om binnen afzienbare tijd de hoofdstukken te publiceren. Daarnaast wil ik je bedanken voor de deur die altijd openstond, waardoor ik alles kon bespreken: van onderzoek tot een vakantie in Zuid-Afrika. Het was fijn om een supervisor te hebben die altijd zo benaderbaar is.

Prof. dr. G. Dijkstra, beste Gerard, de meetings met jou waren altijd erg inspirerend. Vaak kwam ik binnen met een vraag, danwel een probleem en niet zelden ging ik weg met een hoofd vol ideeën. Bedankt voor alle spar- momenten en al jouw out-of-the box ideeën.

Dr. E.A.M. Festen, beste Noortje, ik heb ongelooflijk veel bewondering voor jouw drive; hoe je al jaren het doen van onderzoek en werken in de kliniek weet te combineren. Ik heb de afgelopen jaren veel van je mogen leren. Dank voor alle supervisie- en uitlegmomenten, wanneer de genetica weer eens hogere wiskunde voor me was.

Daarnaast gaat mijn dank uit naar de beoordelingscommissie, Prof. dr. H.W. Verspaget, Prof. dr. H. Snieder en Prof. dr. G.R.A.M. D'Haens, voor de positieve beoordeling van mijn proefschrift.

Tevens wil ik Prof. dr. C. Wijmenga en de afdeling Genetica bedanken voor alle inspirerende meetings en kritische vragen, maar bovenal voor de prettige samenwerking. Bedankt dat ik als IBD-er mocht samenwerken met jouw 'coeliakie-groep'. Kate en Jackie bedankt voor het corrigeren van mijn 'Dunglish'.

Verder wil ik alle mede-auteurs van de artikelen bedanken. Met name Ineke Janse voor haar bijdrage aan hoofdstuk 8 en Thalia Hummel, waarbij het artikel over “medication adherence in adolescents with IBD” dit proefschrift helaas niet heeft gehaald.

IBD-onderzoeksgroep, dank voor alle koffiepauzes, zonder deze momenten was ik nooit zo efficiënt geweest ;). En tevens bedankt dat ik door en met jullie mijn studententijd nog eens drie jaar heb mogen verlengen. Ineke en Valerie, zo groen als gras kwamen jullie drie jaar geleden in onze onderzoeksgroep terecht, maar daar is niets meer van te merken. Fantastisch Val, zoals jij altijd recht voor zijn raap zegt waar het op staat en ik daar het liefst nog een schepje bovenop doe. Tan, ik bewonder je integriteit, maar bovenal moet ik altijd lachen om je droge humor. Marjolein, ik ken jou niet anders dan altijd vrolijk en gezellig. Ik weet wie ik kan bellen als ik weer eens wil stappen. Dianne, door jouw perfectionisme en ambitie weet jij al jaren de Weersma onderzoeksgroep van prachtige data te voorzien. We hadden dit zonder jou niet gekund. Floris en Michiel, ik als niet vegetariër was blij als ik weer eens bij jullie mannenavond mocht aanschuiven voor echte burgers. Michiel, ooit was ik je supervisor, maar onderhand heb je mij allang niet meer nodig. Ik weet zeker dat je je promotie met succes gaat afronden en ik kijk al uit naar het moment dat je ook bij de MDL gaat solliciteren. Floris, el presidente, wat moet ik toch altijd lachen om je grove grappen. Daarnaast heb ik grote bewondering voor je overtuigingskracht. En niet te vergeten, bedankt dat je al mijn frustraties over het onderzoek hebt willen aanhoren, keer op keer. Werna, lieve WTC, ik heb je leren kennen als een ondernemend persoon met een hoofd vol ideeën. Zo onderhand hebben we al tal van uitstapjes ondernomen. Laten we hier vooral niet mee stoppen! Arnau, dank voor al je hulp bij het anti-body stuk. Ik kan me voorstellen dat je mijn hoofd ook wel eens had willen herprogrammeren. Marijn, samen begonnen we aan dit avontuur; ik als jouw studente, jij bezig met de opleiding tot MDL-arts. Gaaf om nu vier jaar later te kunnen zeggen: “dit hebben we toch mooi even gedaan”, jij nu als MDL-arts en ik in opleiding tot MDL-arts. Daarnaast heb ik een enorme bewondering voor je, je bent nooit te beroerd om te helpen en altijd zachtaardig. Ik kijk er naar uit om straks als vrienden naast elkaar te werken.
Dear Boudewijn, Laura, Shixian, Yanni, Ruggero, Ranko, Amber, Ettje, Esther, Joost, Rudi and lab technicians, although not directly involved in my projects, thank you for everything and I wish you all the best the coming years.

San Diego, wat een road trip! Ondanks dat wel elkaar nauwelijks kenden was het een succes. Ik koester de herinneringen aan snickers en sweet sixteen outfits.

Beste MDL (ANIOS, AIOS, staf, secretariaat), mijn affiniteit met de MDL blijkt wel uit dit proefschrift, maar dankzij jullie is mijn enthousiasme, na een jaar op de afdeling MDL te hebben gewerkt, alleen

maar gegroeid. Nu ben ik nog bezig met de vooropleiding interne geneeskunde, maar ik kijk uit naar de dag dat ik weer op de afdeling MDL mag beginnen.

Mirjam Severs, aan enthousiasme geen gebrek bij ons. Waar het begon met een idee voor één project, kwamen er al gauw tig meer. Met eenzelfde werkethos wisten we binnen enkele weken twee (en bijna drie) projecten af te ronden. Ik kon me geen fijnere samenwerking wensen!

Lieve meiden, lieve SpoCo, ondanks dat ik nooit in de SportCommissie heb gezeten, hebben jullie mij geadopteerd alsof het nooit anders is geweest. Lieve Lisanne, twee jaar geleden was het een eer om naast jou als paranimf te mogen staan. Bedankt dat ik al die jaren alle hoogte- en dieptepunten met je heb mogen delen. Het was fijn om hierin een gelijkgestemde te vinden. Lieve Inge, ik vind het fantastisch om te zien hoe jij je ambitie najaagt, al is dat jammer genoeg voor ons wel in het westen. Ik vind je een topper! Lieve Lotte, ik zou willen dat ik net als jij ook dertig uur in een dag kon proppen, wat een multitasker! Veel succes met de huisartsenopleiding. Lieve Gerdia, dank voor je luisterend oor en altijd goede adviezen. Op de een of andere manier weet jij mijn gedachten altijd feilloos te verwoorden. Ik twijfel er geen moment aan dat voor jou later alles op zijn plek valt. Lieve Anne, superleuk dat jij ook in het onderzoek bent gerold, al verbaast het me niks gezien jouw oprechte interesse en enthousiasme, als we het weer eens over ons onderzoek hadden. Heel veel succes met je onderzoek in combinatie met de huisartsenopleiding.
Lieve meiden, jullie vriendschap betekent veel voor mij, hopelijk verandert dit nooit.

Lieve Mariël, ondanks het feit dat we elkaar nog geen jaar kenden, besloten we samen te gaan reizen door Argentinië. Wat hebben we een lol gehad en wat mis ik onze guilty pleasures chemische lolly's met kauwgom en zoute popcorn.

Lieve club, lieve Marjolein, Suzanne, Lysanne, Wieteke en Maritte, ondanks het feit dat jullie niet direct betrokken zijn geweest bij mijn proefschrift, wil ik jullie toch bedanken voor alle gekkigheid, gezelligheid en de super- leuke themafeestjes.

Lieve Nicole en Quirine, dank voor jullie interesse in mijn onderzoek!

Lieve familie Hoorntje, vanaf dag één heb ik me bij jullie thuis gevoeld. Dank voor alle fijne momenten op Schier, waarin jullie me vrij lieten om toch af en toe nog aan mijn proefschrift te werken. Lieve Bien, dank voor het nakijken van mijn Nederlandse samenvatting en dankwoord.

Lieve familie, al kunnen weinig mensen het zich voorstellen, ik kijk elk jaar weer uit naar ons familieweekend met vijf caravans op een boerencamping. Ik hoop dat we onze familietradities zoals Pasen, Kerstmis, Sinterklaas en het familieweekend nog lang mogen voortzetten, het doet mij beseffen hoe belangrijk familie is.

Lieve mam, bedankt voor je luisterend oor de afgelopen jaren en dat ik je altijd kan bellen wanneer dan ook. Fijn als ik weer eens thuis kon komen voor een weekendje schrijven, zonder aan andere dingen te hoeven denken. En sorry dat de kleren, speciaal gekocht voor de promotie, nog een jaar langer in de kast hebben moeten hangen. Lieve pap, met vier vrouwen in huis heb je het vast niet altijd gemakkelijk, toch sta je altijd klaar om te helpen, of nu gaat om de zoveelste verhuizing, of een hor voor het raam. Het enthousiasme voor het programmeren heb ik vast en zeker van jou, het talent helaas wat minder.
Lieve pap en mam, geweldig dat jullie altijd voor mij klaarstaan en me altijd hebben gesteund in mijn keuzes. Bedankt voor alles!

Lieve paranimfen, lieve zusjes, leuk zoals wij altijd kunnen grappen en grollen zonder dat de rest van de wereld er iets van begrijpt. Ik ben blij dat ik de mooie reizen tijdens mijn promotie met jullie heb mogen delen. Lieve Lot, dank voor je hulp bij de vormgeving van dit proefschrift en dank voor alle maaltijden en dan met name de lekkere couscous als ik 's avonds weer eens aan het typen was en geen tijd had om te koken. Lieve Merle, ik bewonder je ondernemerschap en af en toe ben ik erg jaloers op jouw programmeerkunsten.
Lieve zusjes, ik ben ontzettend blij dat jullie op deze speciale dag als paranimfen naast mij staan.

Lieve Edgar, we mogen dan wel heel verschillend zijn, gelukkig sluit onze humor naadloos op elkaar aan. Zonder jouw impulsiviteit en grappen was het maar een saaie boel geweest. Jouw zorgzame en lieve karakter met daarbij eeuwige geduld, hebben er mede toe geleid dat dit proefschrift nu af is! Dank hiervoor.

Curriculum vitae

Lieke Maaike Spekhorst werd geboren op 24 augustus 1988 te Hengelo. Ze haalde in 2007 haar middelbare school diploma aan het lyceum de Grundel in Hengelo en startte aansluitend met de studie Geneeskunde aan de Universiteit van Groningen. Tijdens haar bachelor ging ze naar Tanzania voor een stage in het Namanyere Designated District Hospital. Haar coschappen deed ze in het Universitair Medisch Centrum Groningen in Groningen en in het Medisch Spectrum Twente in Enschede. Tevens was ze tijdens haar studententijd lid van Albertus Magnus en was ze actief in diverse commissies. Na haar wetenschappelijk stage bij de maag-, darm-, en leverziekten, onder begeleiding van Prof. dr. R.K. Weersma, werd ze aangenomen voor het MD/PhD programma in 2013. Nadat zij in 2013 haar studie geneeskunde had afgerond startte zij in 2014 met haar onderzoek. Op het ECCO congres in Barcelona mocht zij de prijs voor 'Top 10 Digital Oral Presentations' in ontvangst nemen. Tevens is Lieke tijdens een deel van haar onderzoek werkzaam geweest in een onderzoekscentrum in Leuven gespecialiseerd in Inflammatory Bowel Diseases (het Translationeel Onderzoek van Gastro-Enterologische Aandoeningen (TARGID)). Na drie jaar onderzoek ging zij in 2017 aan de slag als arts-assistent niet in opleiding op de afdeling maag-, darm-, en leverziekten, in datzelfde jaar werd zij aangenomen voor de opleiding tot maag-, darm-, leverarts. In 2018 wist zij haar promotieonderzoek af te ronden en momenteel is zij werkzaam bij de afdeling algemene interne geneeskunde van het Universitair Medisch Centrum Groningen in het kader van haar opleiding tot maag-, darm-, leverarts.

List of publications

1. F Imhann, J van der Velde, R Barbieri, R Alberts, MD Voskuil, A Vich Vila, V Collij, **LM Spekhorst**, K van der Sloot, V Peters, HM van Dullemen; MC Visschedijk, EAM Festen, M Swertz, G Dijkstra, RK Weersma. The 1000IBD project: multi-omics data of 1000 inflammatory bowel disease patients; Data release 1. *Submitted*

2. Severs M, **Spekhorst LM**, Mangen MJ, Dijkstra G, Löwenberg M, Hoentjen F, van der Meulen-de Jong AE, Pierik M, Ponsioen CY, Bouma G, van der Woude JC, van der Valk ME, Romberg-Camps MJL, Clemens CHM, van de Meeberg P, Mahmmod N, Jansen J, Jharap B, Weersma RK, Oldenburg B, Festen EAM, Fidder HH. Sex-related differences in patients with inflammatory bowel disease; results of two prospective cohort studies. *Inflamm Bowel Dis. 2018 May 18;24(6):1298-1306.*

3. Visschedijk MC, **Spekhorst LM**, Cheng SC, van Loo ES, Jansen BHD, Blokzijl T, Kil H, de Jong DJ, Pierik M, Maljaars JPWJ, van der Woude CJ, van Bodegraven AA, Oldenburg B, Löwenberg M, Nieuwenhuijs VB, Imhann F, van Sommeren S, Alberts R, Xavier RJ, Dijkstra G, Faber KN, Aldaz CM, Weersma RK, Festen EAM. Genomic and expression analyses identify a disease modifying variant for fibrostenotic Crohn's disease. *J Crohns Colitis. 2018 Apr 27;12(5):582-588.*

4. **Spekhorst LM**, Oldenburg B, van Bodegraven AA, de Jong DJ, Imhann F, van der Meulen-de Jong AE, Pierik MJ, van der Woude JC, Dijkstra G, D'Haens G, Löwenberg M, Weersma RK, Festen EAM; Parelsnoer Institute and the Dutch Initiative on Crohn and Colitis. Prevalence of- and risk factors for work disability in Dutch patients with inflammatory bowel disease. *World J Gastroenterol. 2017 Dec 14;23(46):8182-8192.*

5. **Spekhorst LM**, Imhann F, Festen EAM, van Bodegraven AA, de Boer NKH, Bouma G, Fidder HH, d'Haens G, Hoentjen F, Hommes DW, de Jong DJ, Löwenberg M, Maljaars PWJ, van der Meulen-de Jong AE, Oldenburg B, Pierik MJ, Ponsioen CY, Stokkers PC, Verspaget HW, Visschedijk MC, van der Woude CJ, Dijkstra G, Weersma RK; Parelsnoer Institute (PSI) and the Dutch Initiative on Crohn and Colitis (ICC). Cohort profile: design and first results of the Dutch IBD Biobank: a prospective, nationwide biobank of patients with inflammatory bowel disease. *BMJ Open. 2017 Nov 8;7(11):e016695.*

6. **Spekhorst LM**, Severs M, de Boer NKH, Festen EAM, Fidder HH, Hoentjen F, Imhann F, de Jong DJ, van der Meulen-de Jong AE, Pierik MJ, van der Woude CJ, Dijkstra G, Ponsioen CY, Löwenberg M, Oldenburg B, Weersma RK; Parelsnoer Institute and the Dutch Initiative on Crohn and Colitis. The Impact of Ethnicity and Country of Birth on Inflammatory Bowel Disease Phenotype: a Prospective Cohort Study. *J Crohns Colitis. 2017 Dec 4;11(12):1463-1470.*

7. Imhann F, Vich Vila A, Bonder MJ, Fu J, Gevers D, Visschedijk MC, **Spekhorst LM**, Alberts R, Franke L, van Dullemen HM, Ter Steege RWF, Huttenhower C, Dijkstra G, Xavier RJ, Festen EAM, Wijmenga C, Zhernakova A, Weersma RK. Interplay of host genetics and gut microbiota underlying the onset and clinical presentation of inflammatory bowel disease. *Gut. 2018 Jan;67(1):108-119.*

8. Visschedijk MC, Alberts R, Mucha S, Deelen P, de Jong DJ, Pierik M, **Spekhorst LM**, Imhann F, van der Meulen-de Jong AE, van der Woude CJ, van Bodegraven AA, Oldenburg B, Löwenberg M, Dijkstra G, Ellinghaus D, Schreiber S, Wijmenga C; Initiative on Crohn and Colitis; Parelsnoer Institute, Rivas MA, Franke A, van Diemen CC, Weersma RK. Pooled Resequencing of 122 Ulcerative Colitis Genes in a Large Dutch Cohort Suggests Population-Specific Associations of Rare Variants in MUC2. *PLoS One. 2016 Aug 4;11(8):e0159609.*

9. Janse IC, Koldijk MJ, **Spekhorst LM**, Vila AV, Weersma RK, Dijkstra G, Horváth B. Identification of Clinical and Genetic Parameters Associated with Hidradenitis Suppurativa in Inflammatory Bowel Disease. *Inflamm Bowel Dis. 2016 Jan;22(1):106-13.*

10. **Spekhorst LM**, Hummel TZ, Benninga MA, van Rheenen PF, Kindermann A. Adherence to Oral Maintenance Treatment in Adolescents With Inflammatory Bowel Disease. *J Pediatr Gastroenterol Nutr. 2016 Feb;62(2):264-70.*

11. **Spekhorst LM**, Visschedijk MC, Weersma RK, Festen EA. Down the line from genome-wide association studies in inflammatory bowel disease: the resulting clinical benefits and the outlook for the future. *Expert Rev Clin Immunol. 2015 Jan;11(1):33-44.*

12. **Spekhorst LM**, Visschedijk MC, Alberts R, Festen EA, van der Wouden EJ, Dijkstra G, Weersma RK; Dutch Initiative on Crohn and Colitis. Performance of the Montreal classification for inflammatory bowel diseases. *World J Gastroenterol. 2014 Nov 7;20(41):15374-81.*