Personalised medicine in Crohn's disease

Nurulamin M Noor, Bram Verstockt*, Miles Parkes, James C Lee*

Lancet Gastroenterol Hepatol 2020 5: 80–92

*Contributed equally

Department of Medicine, University of Cambridge School of Clinical Medicine, Cambridge Biomedical Campus, Cambridge, UK (N M Noor BMBS, B Verstockt MD, M Parkes DM, J C Lee PhD)**; Department of Gastroenterology, Addenbrooke's Hospital, Cambridge University Hospitals NHS Trust, Cambridge, UK** (N M Noor, M Parkes, J C Lee)**; Medical Research Council Clinical Trials**

Similar to many immune-mediated diseases, Crohn's disease follows a relapsing-remitting pattern, with a variable disease course and heterogeneous clinical outcomes. Frequency of flare-ups, development of complications, and response to treatment collectively determine the effect on a patient's quality of life, which can vary from minimal disruption to profound disability or death. Despite recent advances in the understanding of complex disease pathogenesis, including for Crohn's disease, management decisions are still typically made using a one-size-fits-all approach. Indeed, the inability to reliably predict clinical outcomes in a way that could guide future therapy represents a major unmet need. Recently, several important insights have been made into the biology underlying outcomes in Crohn's disease. In this Review, we will summarise these insights and discuss how greater understanding of these disease mechanisms can be used to develop clinically useful biomarkers, identify novel approaches to optimise disease control, and help deliver the goal of personalised medicine.

Introduction

Crohn's disease typically follows a relapsing-remitting course, but the exact pattern of disease and response to treatment is highly variable between individuals.[1] In recent years, important advances have been made in the understanding of Crohn's disease pathogenesis, in particular through unravelling the genetic architecture of disease susceptibility,[2] the aberrant host immune response,[3] and abnormalities of gut microbiome and metabolome.[4,5] However, despite these advances, the determinants of disease course as well as predictors for response to treatment remain largely unknown.

One of the most important considerations when selecting the treatment approach for individual patients is the disease course or prognosis. For example, a conventional step-up or accelerated step-up approach, in which medications are reactively escalated in response to disease flares will inevitably undertreat patients who are destined to run a more aggressive disease course. Conversely, a top-down approach, in which the most potent therapies are used from the outset (eg, immunomodulators in combination with anti-tumour necrosis factor [TNF] antibodies) has been shown to improve clinical outcomes in Crohn's disease,[6,7] but would expose patients destined to have mild disease to unnecessary risks. Moreover, such a strategy would be unaffordable in most healthcare settings around the world. Indeed, the variability in the course of Crohn's disease means that any one-size-fits-all approach will be suboptimal for a substantial proportion of patients. There is now a widespread consensus that a more individualised approach is required, with a need for accurate biomarkers that can enable the right patient to be matched to the right treatment (figure 1, panel). This individualised approach is not only an unmet need in Crohn's disease, but also in most autoimmune and inflammatory conditions.[8]

Although the concept of personalised medicine in inflammatory bowel disease (IBD) might seem unrealistic, given that this idea has been discussed for years and yet seems no closer to clinical practice, it is clear that this concept has already been achieved in other fields—notably, oncology. Indeed, one of the first and most commonly cited applications of a personalised approach to treatment is the example of breast cancer, in which patients with tumours expressing HER2 have been shown to specifically benefit from treatment with

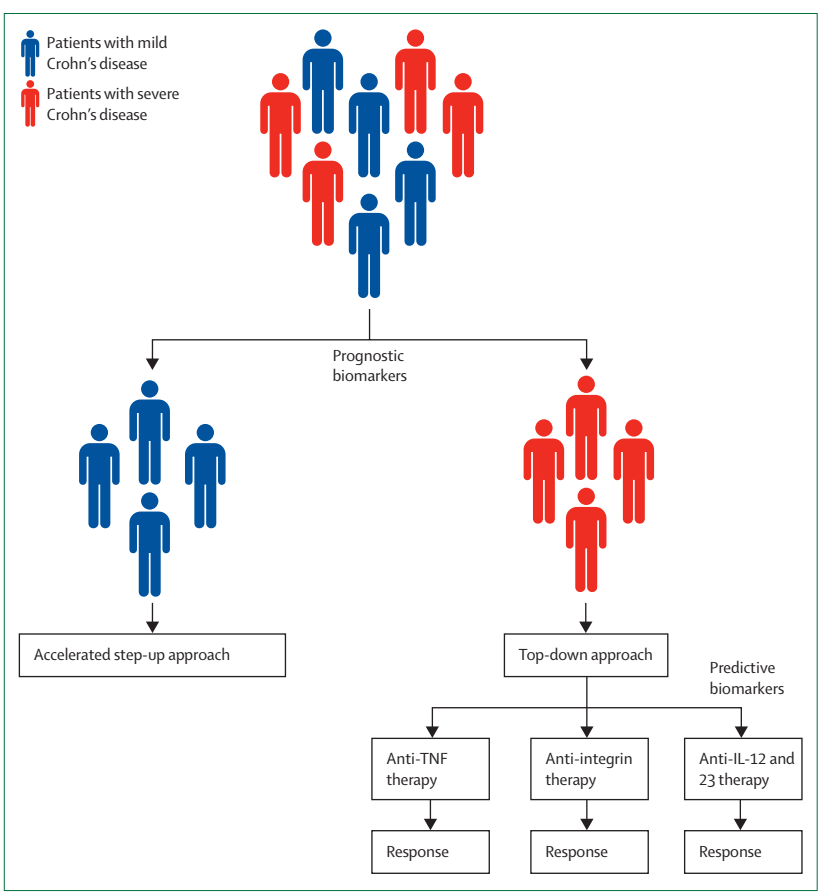

Figure 1: **Clinical use of biomarkers to direct a personalised approach for Crohn's disease**
An exemplar of how personalised medicine could be delivered by incorporating prognostic and predictive biomarkers into clinical decision making to enable the most appropriate treatment for an individual to be selected. Prognostic biomarkers could identify at the time of diagnosis which patients will have more aggressive disease and so would require more potent therapy and predictive biomarkers could then be used to match patients to the most appropriate therapy. This approach would avoid risks of unnecessary immunosuppresion in patients with more indolent disease, and ensure that patients with aggressive disease receive the most appropriate therapy as early as possible. TNF=tumour necrosis factor. IL=interleukin.

trastuzumab therapy, a monoclonal antibody that targets HER2.[9] Conversely, those with tumours that did not express this receptor displayed worse clinical outcomes when given the same treatment.[9] Subsequently, more contemporary approaches have been applied to breast cancer, one notable example being the application of transcriptional profiling to predict prognosis and help stratify choice of therapy for patients.[10]

In Crohn's disease, and in inflammatory diseases as a whole, progress has been more limited. One reason for this slow progress has been the difficultly in determining which tissues to study since there is no equivalent to the tumour in oncology. In Crohn's disease, this debate has included discussion of whether to use specific cell types or mixed tissues such as whole blood or intestinal biopsies, and how to analyse data to ensure accurate assessment of the constituent cell types.

Recently there has been an increasing drive towards personalised or precision medicine across all fields, spurred on by large-scale initiatives designed to support such research efforts, including the Precision Medicine Initiative in the USA,[11] and the Pharmacogenetics and Stratified Medicine Network in the UK.[12] The encouragement from the scientific community and associated funding calls will have a key role in driving further developments in personalised medicine.

In this Review, we aim to discuss how understanding the mechanisms of Crohn's disease can be used to develop clinically useful biomarkers, identify novel approaches to optimise disease control, and help deliver the goal of personalised medicine.

Predicting disease outcome and prognosis

Clinical predictors of disease outcome

Several clinical predictors of a poor prognosis have long been proposed in Crohn's disease, including the presence of peri-anal disease, ileocolonic disease, upper gastrointestinal involvement, a diagnosis before the age of 40 years, and a requirement for steroid treatment at initial presentation.[13] More recently, other factors have also been suggested as predictors of poor prognosis, including smoking and severe endoscopic appearances.[14] Conversely, several clinical factors have been proposed as predictors of a better prognosis, including not smoking, achieving a higher educational level, rectal sparing (ie, disease that does not involve the rectum), and older age at diagnosis.[15] However, all of these predictors have been derived from observational studies, usually done retrospectively, and are unable to reliably distinguish patient subgroups in a way that can correctly guide treatment decisions for the majority of patients with Crohn's disease. Indeed, it is clear that the distinction between association and prediction, or rather the failure to appreciate the differences in these parameters, has led to many weak associations being reported as so-called predictors, despite these factors not showing predictive performance even in the same discovery cohorts. Accordingly, improved prediction of disease course remains one of the major unmet needs in Crohn's disease management and a research priority in the field.[8] The ideal prognostic biomarker would be one which could be tested at diagnosis, or as close to diagnosis as feasible, to maximise the window of opportunity that exists early in the disease before clinically significant bowel damage might have occurred. Indeed, the value of a prognostic biomarker diminishes the further the patient is from diagnosis, as the extent of their Crohn's disease is likely to have already become apparent.

Panel: Summary of biomarkers that have been validated or are undergoing validation for clinical application in Crohn's disease

Prognostic biomarkers for use at diagnosis
- PredictSURE IBD

Prediction of adverse effects
- *TPMT* variants and thiopurine-induced myelosuppression
- *NUDT15* variants and thiopurine-induced myelosuppression
- HLA-DQA1-HLA-DRB1 variants and thiopurine-induced pancreatitis

Prediction of response or non-response

Anti-TNF
- HLA-DQA1*05 variants and immunogenicity
- *OSM* gene expression and response
- TREM-1 and response

Anti-integrin
- αE gene expression (*ITGAE*) and response

Anti-IL-12 and 23

Serum IL-22 and response

IBD=inflammatory bowel disease. TNF=tumour necrosis factor. IL=interleukin.

Genetic predictors of disease outcome

Following the enormous success of genetic studies in identifying single nucleotide polymorphisms (SNPs) that are associated with the development of Crohn's disease,[2,16] it was no surprise that many investigators subsequently sought to identify genetic predictors of prognosis (table 1). Similar to disease susceptibility studies, however, it is important to note that almost all of this work has been done in European populations—mainly because of the availability of patient cohorts—and as such might not generalise to other populations. Early examples of such studies investigated whether known susceptibility SNPs might also influence clinical outcomes. For example, several studies have reported associations between Crohn's disease-associated variants in the *NOD2* gene with need for surgery and fibrostenotic disease.[23] *NOD2* gene variants were subsequently included in composite risk models (in conjunction with clinical and serological variables) for both paediatric

Unit, University College London, London, UK (N M Noor); **Department of Gastroenterology and Hepatology, University Hospitals Leuven, Katholieke Universiteit Leuven, Leuven, Belgium** (B Verstockt); **Translational Research in Gastrointestinal Disorders (TARGID), Department of Chronic Diseases, Metabolism and Ageing, Katholieke Universiteit Leuven, Leuven, Belgium** (B Verstockt); **and Cambridge Institute for Therapeutic Immunology and Infectious Disease, Jeffrey Cheah Building, Cambridge Biomedical Campus, Cambridge, UK** (J C Lee)

Correspondence to:
Dr James C Lee, Department of Medicine, University of Cambridge School of Clinical Medicine, Cambridge Biomedical Campus, Cambridge, CB2 0QQ, UK
jcl65@cam.ac.uk

	Chromosome	SNP	p value	OR (95% CI)	AF	Candidate gene in or near locus
Disease location						
Cleynen et al (2016)[17]	6	rs6930777	$8{\cdot}13\times10^{-23}$	0·68 (0·62–0·75)	0·11	MHC
Cleynen et al (2016)[17]	16	rs2066847	$1{\cdot}01\times10^{-35}$	1·89 (1·62–2·21)	0·02	*NOD2*
Disease behaviour						
Cleynen et al (2016)[17]	6	rs77005575	$2{\cdot}82\times10^{-10}$	1·16 (1·12–1·21)	0·44	MHC
Cleynen et al (2016)[17]	16	rs2066847	$5{\cdot}73\times10^{-10}$	1·31 (1·21–1·42)	0·02	*NOD2*
Prognosis						
Lee et al (2017)[18]	X	rs5929166	$4{\cdot}56\times10^{-9}$	0·33 (0·23–0·48)	0·03	*XACT*
Lee et al (2017)[18]	6	rs9279411	$5{\cdot}46\times10^{-9}$	0·60 (0·50–0·71)	0·15	MHC
Lee et al (2017)[18]	6	rs147856773	$1{\cdot}31\times10^{-8}$	0·57 (0·47–0·70)	0·12	*FOXO3*
Lee et al (2017)[18]	7	rs75764599	$4{\cdot}32\times10^{-8}$	3·02 (2·04–4·49)	0·01	*IGFBP1-IGFBP3*
Treatment-associated adverse effects						
Thiopurine-induced myelosuppression						
Walker et al (2019)[19]	6	rs11969064	$5{\cdot}2\times10^{-9}$	2·3 (1·7–3·1)	0·09	*TPMT*
Yang et al (2014)[20]	13	rs116855232	$4{\cdot}9\times10^{-94}$*	35·6 (22·5–56·5)*	0·55*	*NUDT15*
Walker et al (2019)[19]	13	rs746071566	$1{\cdot}3\times10^{-8}$†	38·2 (5·1–286·1)†	<0·01†	*NUDT15*
Thiopurine-induced pancreatitis						
Heap et al (2014)[21]	6	HLA-DQA1-HLA-DRB1 (rs2647087)	$2{\cdot}0\times10^{-16}$	2·59 (2·1–3·3)	0·28	MHC
Anti-TNF immunogenicity						
Sazonovs et al (2018)[22]	6	HLA-DQA1*05 (rs2097432)	$5{\cdot}9\times10^{-13}$	1·9 (1·6–2·3)	0·22	MHC

AF=allele frequency. OR=odds ratio. MHC=major histocompatibility complex. SNP=single nucleotide polymorphism. TNF=tumour necrosis factor. *Data from a South Korean cohort. †Data from a cohort with European ancestry.

Table 1: **Genetic variants associated with different aspects of Crohn's disease**

Crohn's disease[24] and adult Crohn's disease to form a web-based personalised risk and outcome prediction tool.[25] *NOD2* variants were also included in a study of progression to surgery in 1115 genotyped participants.[26] Patients identified as higher risk based on genotype and clinical risk factors had earlier progression to surgery than those identified by clinical risk factors alone. However, an important confounder for these prediction models is that *NOD2* variants are known to be specifically associated with ileal Crohn's disease, for which surgery is more commonly performed, as the operation has a lower morbidity and mortality compared with other surgeries (eg, colectomy).[27] Indeed, it was subsequently shown that when disease location is taken into account, the link between *NOD2* and the need for surgery disappears.[17] Unsurprisingly, other loci that also associate with ileal disease, such as *ATG16L1* and *IRGM*, have similarly been reported to be associated with disease course, but this association is not replicated when disease location is taken into account.[28]

Following the initial success of genome-wide association studies (GWASs), it was hypothesised in several disease areas that disease course might relate to the overall burden of disease susceptibility variants, so that individuals with more risk alleles would have more severe forms of disease. This hypothesis would be consistent with a model in which there is a genetic continuum from health to mild disease and then on to more severe disease phenotypes, with susceptibility variants also acting as modifiers of disease course.[29] Indeed, such a relationship between the number of risk variants and phenotypic severity is observed for quantitative traits such as height. However, when applied to Crohn's disease using a genetic risk score incorporating strength of association and allele dosage, risk loci could only explain a tiny fraction of variance of disease phenotype.[30]

Because of this inability to identify robust associations between susceptibility variants and disease course, focus has subsequently turned to non-susceptibility variants as potential determinants of outcome—ie, genes or variants, or both, that are not involved in conferring disease risk. Accordingly, a GWAS of prognosis in Crohn's disease, the first in any inflammatory disease, was done in two cohorts of European patients comparing those who had an aggressive disease course with those who had much more indolent disease (as assessed by the treatments they had been given).[18] Meta-analysis of the combined dataset of 2734 patients was able to identify four distinct loci that were associated with prognosis (*FOXO3*, *XACT*, *IGFBP1*, and the major histocompatibility complex region stretching from the *HLA-B* to *HLA-DR* genes). Strikingly, none of these four prognosis-associated

variants had previously been associated with susceptibility and neither the aggregate effect of all susceptibility variants for Crohn's disease nor any individual susceptibility variants showed a meaningful association with prognosis. This finding provides increasing evidence for there being distinct genetic and biological underpinnings for susceptibility and prognosis, as well as disproving the previously held notion that an increasing number of risk alleles would lead to a more severe disease course.[29] Such studies have helped to provide important clues about the underlying biology of prognosis in Crohn's disease, as shown by the *FOXO3* genetic variant, which had previously been identified in a candidate gene study of prognosis in Crohn's disease and shown to regulate inflammatory responses in monocytes, via a TGFβ1-dependent pathway.[31] Although the results of such studies are able to provide important insights into disease biology, it is noteworthy that these are unlikely to form the basis of a prognostic test, since the low odds ratios (ORs) observed for individual prognosis-associated variants, means that any such test would not be sufficiently discriminatory. This effect mirrors the situation in susceptibility GWAS in which the results can provide useful insights into biology but cannot predict who will develop disease.

The one area of genetic research where ORs for SNP associations are large enough to allow predictive testing, however, is the field of pharmacogenetics. For example, it is already common practice to assess thiopurine methyltransferase activity or genotype status before starting a patient on thiopurine medications. More recently, a non-synonymous SNP in the *NUDT15* gene has been associated with thiopurine-induced myelosuppression (initial OR >35) reported for a cohort of South Korean patients.[20] A subsequent exome-wide association study with replication in an independent validation cohort, has confirmed that carriage of any of the three coding variants of *NUDT15* was associated with thiopurine-induced myelosuppression in patients of European ancestry (OR >27).[19] Similarly, genetic associations have also been reported with thiopurine-induced pancreatitis,[21] 5-aminosalicylate-induced nephrotoxicity,[32] and anti-drug antibody formation in anti-TNF-treated patients.[33] In the future, one could certainly envisage using a panel of SNP-based genetic tests in patients with newly diagnosed Crohn's disease to help guide appropriate treatment selection.

Transcriptomic predictors of disease outcome

By studying gene expression, the RISK study (NCT00790543) has provided recent insights into Crohn's disease prognosis.[34] Transcriptomic analysis of ileal biopsies from 359 newly diagnosed, treatment-naive paediatric patients with IBD (n=243 with Crohn's disease) was used create a model combining clinical, transcriptional, and microbial data, which showed superiority to a model containing only clinical phenotyping data.[35] Furthermore, an extracellular matrix signature obtained from ileal biopsies in an extended cohort of patients from the RISK study (n=913 with Crohn's disease) showed an ability to predict later development of stricturing disease, when followed up prospectively for 36 months.[34] Work is currently ongoing to further refine this model and develop a risk score for more severe disease, using integrated data from RNA sequencing along with expression quantitative trait loci from peripheral blood—whereby genetic variation affects transcription of RNA.[36] This work emphasises the value of prospectively collected cohorts sampled before receiving therapy, but does have some limitations. First, because the patients in the RISK cohort were assessed and treated at the discretion of their treating physician, rather than following a formal protocol, conclusions about the possible effects of treatments on disease course should be made with caution. Second, the performance of the prognostic markers identified was assessed in the same cohort of patients in which they were discovered (using split sample validation or leave one out cross-validation). These methods involve resampling the same cohort to estimate the performance of the detected biomarkers, but in the absence of a separate replication cohort, it is hard to know whether this predictive performance would be generalisable to other patient cohorts. There have been other examples of biomarkers that have been too tightly modelled on the training dataset, and so do not work when tested in new cohorts.[37] Therefore, to determine whether the prognostic markers from the RISK cohort are broadly applicable to wider populations, any such model ideally needs to be examined in an independent validation cohort.

Because of the difficulty in distinguishing the relative cellular components within a mixed cell population in intestinal samples,[38] and the inherent requirement for an endoscopic procedure, other studies have examined whether clinically useful biomarkers could be developed from whole blood. Early work in this area identified a gene expression signature in isolated CD8+ T cells that was shown to strongly correlate with the tendency to flare early and recurrently for both Crohn's disease and ulcerative colitis. Interestingly, this finding was analogous to a prognostic CD8+ T-cell signature that had previously been described in anti-neutrophil cytoplasmic antibody-associated vasculitis and systemic lupus erythematosus, which are also associated with prognosis.[39] Across all of these immune-mediated diseases, the gene expression signature divided all patients into two distinct subgroups, which were clinically indistinguishable at presentation but had very different courses of disease thereafter. Subsequent exploration of the biology responsible for these transcriptional differences showed that patients in the good prognosis subgroup were enriched for a signature of CD8+ T-cell exhaustion—a phenomenon in which antigen-experienced T cells progressively lose their ability to respond to their target

antigen.[40] An important role for this cell subtype is also supported by work demonstrating that a substantial increase in the number of CD8+ T-cell clones is associated with worse outcomes and relapse for patients with Crohn's disease in the postoperative period.[41]

The aforementioned CD8+ T-cell biomarker has since been developed into a multi-gene, quantitative polymerase chain reaction (qPCR) assay that can be done on unseparated whole blood, immediately widening the translational potential of such a test, and making it accessible to standard hospital laboratories around the world.[42] Following its development using machine learning, this genomic biomarker was validated in independent cohorts of newly diagnosed patients with Crohn's disease or ulcerative colitis from the UK. Moreover, this whole-blood qPCR biomarker is now being assessed for its ability to guide treatment from diagnosis in a UK trial of 400 patients with newly diagnosed Crohn's disease (ISRCTN11808228).[43] This trial is the first biomarker-stratified study in IBD, and will directly assess whether personalised medicine can be delivered to patients with Crohn's disease.

Microbial predictors of disease outcome

Several studies have reported that circulating antibodies against bacterial antigens (including anti-I2, anti-ompC, ASCA, pANCA, anti-CBir1 flagellin) are associated with a more complex Crohn's disease phenotype and are more common in patients with a stricturing or penetrating phenotype.[44] Similarly, meta-analyses have shown an association between ASCA-positivity and the need for surgery.[45] On the basis of such studies, antibody panels have been developed for prognostic tests in IBD, but it remains unclear whether these antibodies are present as a cause or effect of more severe disease, since it is widely reported that seropositivity increases with duration of disease.[46] As a result, and because of the limited predictive value of these markers, they are currently not recommended to help guide management of Crohn's disease in routine clinical practice.[47]

Other studies have sought to correlate the clinical course of Crohn's disease with specific species of bacteria in the intestine. Caution needs to be exercised when interpreting the results of such studies, since microbial changes are known to occur because of a broad range of potential confounders including inflammation, diet, and smoking.[48] In spite of this limitation, there have been some promising reports including the demonstration that a reduced carriage of *Faecalibacterium prausnitzii*, a butyrate-producing *Clostridium* species, is associated with endoscopic recurrence at 6 months after resection.[49] Whether this relationship is causal (ie, dysbiosis causes inflammation) or consequential (ie, inflammation causes dysbiosis) remains to be fully clarified. There is certainly more evidence for a consequential relationship, although a recent study showed that transferring intestinal microbiota from patients with IBD exacerbated a mouse model of colitis, suggesting that dysbiosis might not only be a result of inflammation.[50] Further work to resolve these relative effects is clearly necessary, particularly if microbiota assessment is ever to be considered as a prognostic tool in clinical practice.

Moreover, substantial changes in microbial diversity can be attributed to environmental factors, such as smoking,[51] and dietary influences.[52] In this regard, the PREdiCCt study (ISRCTN11808228), and other similar environmental studies are important, as they will combine all these environmental elements with genetic risk and detailed clinical phenotyping data. Such studies should provide novel insights and potentially highlight key drivers behind both relapse and disease remission in Crohn's disease.

Proteomic predictors of disease outcome

C-reactive protein is an established marker of disease activity in Crohn's disease and is routinely used for measuring inflammation. C-reactive protein concentrations of more than 10 mg/L 1 year after diagnosis have been shown to be predictive for progression to abdominal surgery in a prospective Norwegian population-based study of patients with Crohn's disease.[53] Similarly, for patients in clinical remission with persistently elevated concentrations, an inverse correlation has been shown with long-term outcome as measured by the number of hospital admissions and intestinal resections.[54] However, it is important to recognise that C-reactive protein is a marker of current, ongoing inflammation, and thus the described associations are likely to simply reflect the fact that ongoing, inadequately treated inflammation leads to more complications than occur once disease is in remission.

The CALM trial has shown that normalising faecal calprotectin can be a useful treatment target when combined with other parameters, and is associated with mucosal healing in Crohn's disease.[55] Calprotectin is well described as a validated marker for disease activity but has also been reported to be a useful biomarker for predicting relapse. In ulcerative colitis, a faecal calprotectin concentration of less than 50 µg/g was able to predict endoscopic remission at week 10 following treatment with infliximab,[56] whereas an elevated calprotectin concentration of more than 300 µg/g correlated with higher likelihood of relapse, despite no clinical symptoms to suggest a flare of disease.[57] In Crohn's disease, the POCER study[58] measured faecal calprotectin repeatedly in 104 patients following ileocaecal resection, and reported that a faecal calprotectin concentration of more than 100 µg/g was positively associated with endoscopic recurrence (Rutgeerts score ≥i2), with a negative predictive value of 91%. Notwithstanding the relatively low bar for endoscopic recurrence, this finding suggests that a normal faecal calprotectin concentration in postoperative patients could indicate that colonoscopy is unnecessary.

Important limitations of using faecal calprotectin concentration **as a** predictor of disease outcome exist. These limitations include variability in the faecal calprotectin result, which has been shown both for commercially available assays, and for samples taken from the same patient on the same day using the same assay, and an increasing awareness of the effect of bowel movement frequency and stool consistency on the result.[59] There are also logistical challenges of using calprotectin in Crohn's disease—relating to both the low return rate of calprotectin samples from real-world clinical cohorts and more specifically to lower concentrations of faecal calprotectin in patients with small bowel inflammation.[60] Despite the recent advances of proteomic applications to Crohn's disease, few studies have focused on disease outcome. Two such studies, which will hopefully offer novel and important contributions, are the IBD Character Study[61] (a multicentre European project) and the PREDICTS study across North America,[62] both of which are seeking to identify prognostic proteomic biomarkers.

Predicting treatment response and non-response

Currently, the best predictor of response or non-response to a therapy (including thiopurines, anti-TNFα, and anti-integrins) is based on confirming adequate drug exposure. This approach has been well outlined for the use of thiopurine metabolites to optimise outcomes,[63] and increasingly for newer agents in Crohn's disease including therapeutic drug monitoring of anti-TNF therapy.[64] Indeed, in the PANTS study, the only factor that was independently associated with primary non-response to anti-TNFα therapy was a low drug concentration at week 14.[33] Although this strategy can help improve treatment response in patients with low drug concentrations, there remains a clear unmet need for prediction tools that can be used before treatment is commenced to help match patients to the agent to which they have the best chance of responding. It is also noteworthy that prospective infliximab dose optimisation, based on therapeutic drug monitoring, was recently shown not to improve clinical outcomes in a cohort of 122 patients with Crohn's disease.[65]

Anti-TNF

Targeting TNFα has resulted in substantial improvements in clinical care in Crohn's disease.[6] However, inadequate response to initial treatment with anti-TNFα therapy, has been reported to be as high as 40%.[64]

Clinical predictors of response

Some clinical features have been reported to be associated with primary non-response to anti-TNF therapy, including disease duration for more than 2 years[66] and ongoing smoking.[67] However there are also several contradictory reports of clinical features that are associated with reduced response rates, including both a normal C-reactive protein concentration[68] and an elevated C-reactive protein concentration before anti-TNF therapy,[69] and both isolated small bowel disease[70] and isolated colonic disease.[71] This lack of congruity is perhaps not surprising, as many of these associations have been retrospectively identified in small cohorts of patients and not externally validated. As such, these observations are currently insufficient to guide clinical decision making.

Mucosal immune predictors of response

An elegant translational application was shown in 25 patients with active Crohn's disease, defined by a Crohn's disease activity index of 150 or more, who underwent endoscopic examination before initiation of adalimumab treatment.[72] Adalimumab was fluorescently labelled and applied to the most inflamed areas of the bowel using a spray catheter. Confocal imaging—to enumerate the cells with membrane-bound TNFα—allowed patients to be stratified into two groups using a cutoff of 20 TNFα-positive cells per confocal image. Patients with fewer positive cells had a 15% clinical response to adalimumab at 12 weeks, whereas those with higher number of positive cells had a 92% response. Although this finding needs to be validated in an external cohort, the concept of combining endoscopy, imaging, and immunology to predict anti-TNFα response is promising for practising gastroenterologists.

Genetic predictors of response

Genetic markers of clinical response to infliximab, adalimumab, or certolizumab at week 12 were examined in 359 genotyped patients with Crohn's disease.[73] A combined risk score of clinical covariates (age, sex, smoking status, location and phenotypes of Crohn's disease, and presence of perianal involvement) and a genetic risk score (including 15 identified risk alleles) was shown to be superior to a model only containing the clinical covariates for predicting non-response. However, the previously discussed issue of low ORs from observed genetic associations might limit the ability to develop a robust prediction tool using SNPs. The initial optimism for genetic approaches has been tempered by more recent genome-wide applications, which have reported that disease susceptibility loci do not contribute to anti-TNF non-response,[74] and similarly that polymorphisms within the *TNF* gene or the TNF-receptor pathway do not associate with treatment response.[68]

However, findings from the PANTS study,[33] a large UK-wide, prospective, observational study, which has enrolled more than 955 patients, show promise for genetic applications. The study examined response to anti-TNF therapy and provided important insights into the genetics of anti-drug antibody formation.[33] Specifically, the authors showed that formation of anti-drug antibodies appears to be significantly enriched in 40% of patients of European ancestry carrying the HLA-DQA1*05 variant. Although the carriage of this variant will be important to determine

across populations of non-European ancestry, pre-treatment genetic testing for this variant might prove potentially practice-changing in order to appropriately help stratify patients and maximise response to anti-TNF therapy.[22]

Transcriptomic predictors of response

Initial transcriptomic studies aiming to identify predictors of treatment response were done using intestinal biopsies from patients with ulcerative colitis. In a study of 45 patients from two cohorts, a gene expression signature was described that correlated with endoscopic response to infliximab at 4–8 weeks.[75] Applying the same technology to Crohn's disease, a study using intestinal biopsies from 37 patients (19 with colonic Crohn's disease, nine with ileocolonic Crohn's disease, and nine with small bowel involvement only) aimed to predict response to infliximab at 4–6 weeks, and showed substantial overlap of predictive genes between patients with Crohn's disease and those with ulcerative colitis.[76] The top five genes were able to predict response to infliximab induction in colonic Crohn's disease with 100% accuracy, although this finding was validated in the same cohort in which it was discovered and so it remains unclear whether this effect will be generalisable. Interestingly, this approach was unable to identify a predictive panel for determining response or non-response in ileal or ileocolonic disease, highlighting the importance of further understanding molecular sub-phenotyping in Crohn's disease.[17,77]

An important methodological point—particularly for studies using data in which the number of parameters measured far exceeds the number of samples, such as transcriptomics—is that external validation is necessary to be confident that any reported associations have not occurred by chance. This idea has been recently illustrated by the attempt to refine and optimise the above-described multi-gene transcriptomic biomarker for anti-TNF response in ulcerative colitis. This biomarker was initially refined using biopsy samples from the PURSUIT trial (NCT00488631) of golimumab therapy, and then externally tested in a separate cohort using samples from the PROgECT phase 2a trial (NCT01988961).[37] Unfortunately, despite excellent performance in the cohorts in which it had been developed, this biomarker did not show sufficient predictive ability in an independent cohort, highlighting that the performance of a biomarker in new cohorts cannot be assumed to be the same as in the discovery cohort without first being formally tested. The PROgECT trial does, however, emphasise the need for combining expertise and efforts across industry and academia to help develop clinically useful biomarkers in the future.

One such academia–industry collaboration is the promising application of mucosal transcriptomic technology from five datasets comprising 227 patients with either Crohn's disease or ulcerative colitis.[78] Higher expression of Oncostatin M (OSM), a member of the IL-6 pro-inflammatory cytokine family, was associated with greater disease activity and reported to be predictive of non-response to anti-TNF therapy. It remains unclear whether OSM also represents a tractable therapeutic target or is simply a marker for non-TNFα inflammation, which would be consistent with previous work highlighting that primary non-response might be due to a high inflammatory burden.[79] The OSM discovery study highlights the strength of, and need for, academia–industry collaboration, as initial transcriptomic data were validated from multiple, publicly available datasets as well as unpublished transcriptomic data from the ACT (NCT00207688) and PURSUIT trials.

Publicly available datasets have also facilitated the identification of other potential biomarkers. For example, using cell-based deconvolution methods, low levels of TREM-1 has been reported to predict anti-TNF non-response,[80] as has CCR2-CCL7 axis upregulation—potentially reflecting an increase in inflammatory macrophages.[81] Translating such observations to clinical practice remains a challenge and there have been contrasting findings on whether biomarkers such as TREM-1 can be measured in whole blood and still retain predictive performance.[82] It is clear, however, that delineation of data from multiple cell types should offer insights into disease biology and will become increasingly more commonplace given the rapid advances of single-cell technologies. Indeed, single-cell analysis of ileal Crohn's disease lesions has already been reported to provide a biomarker of anti-TNFα response in a study comprising four independent cohorts of patients with Crohn's disease.[83] The downstream work from these approaches will probably pave the way for future biomarker discovery and tailored therapeutic opportunities.

Microbial predictors of response

Although work is ongoing to discern the cause–effect relationship between inflammation and microbial dysbiosis, more recent studies have sought to investigate how therapy might affect the gut microbiome and whether there might be changes in community structure or metabolites, or both, that could correlate with treatment response. For example, a small but prospective study of patients receiving anti-TNFα recently correlated predicted concentrations of butyrate with response to therapy.[5] An earlier study in paediatric IBD also correlated anti-TNF response with microbial composition—with higher responses in those with gut microbiota that more closely resembled healthy individuals without IBD.[84] As with other microbiome studies in IBD, this finding might simply reflect a lower inflammatory burden and has not been replicated in all studies,[5] but certainly warrants further investigation and highlights the potential for using the gut microbiome as a predictive tool.

Anti-integrin

Vedolizumab, an antibody targeting the α4β7 heterodimeric integrin receptor, has proven to be an effective treatment in patients with Crohn's disease to induce and maintain clinical remission.[85]

Clinical predictors of response

A common finding reported across late-phase, randomised investigations of novel biologic agents in Crohn's disease, is the greater response shown in patients naive to anti-TNF compared with anti-TNF exposed patients, a finding particularly notable with vedolizumab.[86] Subsequently, a scoring system has been developed from the GEMINI trials to identify patients with Crohn's disease most likely to respond to vedolizumab at 26 weeks and validated in a cohort of 366 patients from the VICTORY consortium.[87] This clinical prediction score is composed of previous treatment with anti-TNF therapy, bowel surgery, fistulating disease, as well as baseline albumin and C-reactive protein concentrations. However, these features are recognised signs of more severe disease, and so an association with non-response is perhaps unsurprising. Moreover, such approaches are also subject to the limitations discussed previously, and there remains a need for prospective identification of potential prognostic factors and then appropriate and robust validation of these features in independent cohorts of patients.

Mucosal predictors of response

Similar to anti-TNF therapy, fluorescently tagged antibodies have been used to study prediction of response to vedolizumab treatment.[88] By targeting the α4β7 integrin, vedolizumab prevents binding to MAdCAM-1, thereby inhibiting movement of α4β7-expressing cells (predominantly lymphocytes and monocytes) to lymphoid tissues in the gut and hence reducing the mediation of inflammation. In a small proof-of-principle study, involving five patients with Crohn's disease, α4β7-expressing mucosal cells in pericryptal regions of the gut, obtained following intestinal biopsy, were labelled ex vivo with fluorescein isothiocyanate.[88] Clinical and macroscopic endoscopic responses were reported in the two patients with α4β7-expressing mucosal cells, whereas non-response was shown in those not expressing these cells. Translation of this work is ongoing but if this process could be developed in vivo and applied to allow real-time use, then one could envisage such technology being used as an adjunct during future endoscopic procedures.

Indeed, measuring the amount of molecule that is targeted by a particular therapy to guide any potential therapies and their likely efficacy warrants further discussion. It could be predicted, for example, that in the future endoscopic techniques will not only be used for making a diagnosis and assessing disease activity, but also to help guide therapy by assessing the relative abundance of target cells. Likewise, biopsy samples could undergo processing and transcriptomic analysis for markers such as *OSM* and *TREM1*, and the resulting information could be used to allow selection of the most appropriate therapy for that individual patient. This practice would be akin to those in oncology in which analysis of tumours can facilitate informative discussions about most appropriate therapeutic options.

The question then arises about whether future biomarkers in Crohn's disease should be based on blood or biopsy. In the early stages of biomarker discovery, biopsy-based biomarkers were more widely studied because they were considered to capture the most relevant immune responses. However, it is also clear that most immune cells circulate between the gut and blood and so can be sampled in either. Indeed, given the relative simplicity of a blood-draw, it is probable that a blood-based biomarker would be more acceptable to most patients and gastroenterologists. Ultimately, however, the best markers will be ones that can most effectively guide clinical decision making and improve outcomes for patients, irrespective of the origin of the tissue.

Genetic and transcriptomic predictors of response

Predicting response to etrolizumab, which binds to the β7 subunit of α4β7 and αEβ7 integrin heterodimers, has again focused mostly on transcriptomic investigation of mucosal biopsies, with most studies to date focusing on ulcerative colitis. In the EUCALYPTUS phase 2 trial (NCT02100696) of 119 patients with ulcerative colitis treated with etrolizumab, αE gene expression (*ITGAE*) measured by RT-qPCR was able to stratify those with greater likelihood of clinical remission at week 10.[89] Further exploration of this mechanism suggested an association between αE-expressing cells and the presence of granzyme A, a serine protease present in activated and cytotoxic T cells and natural killer cells.[90]

Further transcriptomic analysis of mucosal biopsies in 41 patients with ulcerative colitis has shown substantial changes in gene probe sets predictive of endoscopic response to vedolizumab at week 52.[91] Although similar findings need to be shown in Crohn's disease, 63% of the genes that predicted response to vedolizumab also predicted response to infliximab.[91] These findings suggest that response might not be treatment specific and that predictors of response or non-response might apply to a range of biologic medications, as they all target the common pathways of inflammation.

Microbial predictors of response

A microbial approach to predict response to vedolizumab has included the development of a network algorithm, vedoNet, which uses a combination of microbiome and clinical phenotyping data.[92] 85 patients with IBD (43 with ulcerative colitis and 42 with Crohn's disease) with a mean disease duration of 13 years were enrolled before initiation of

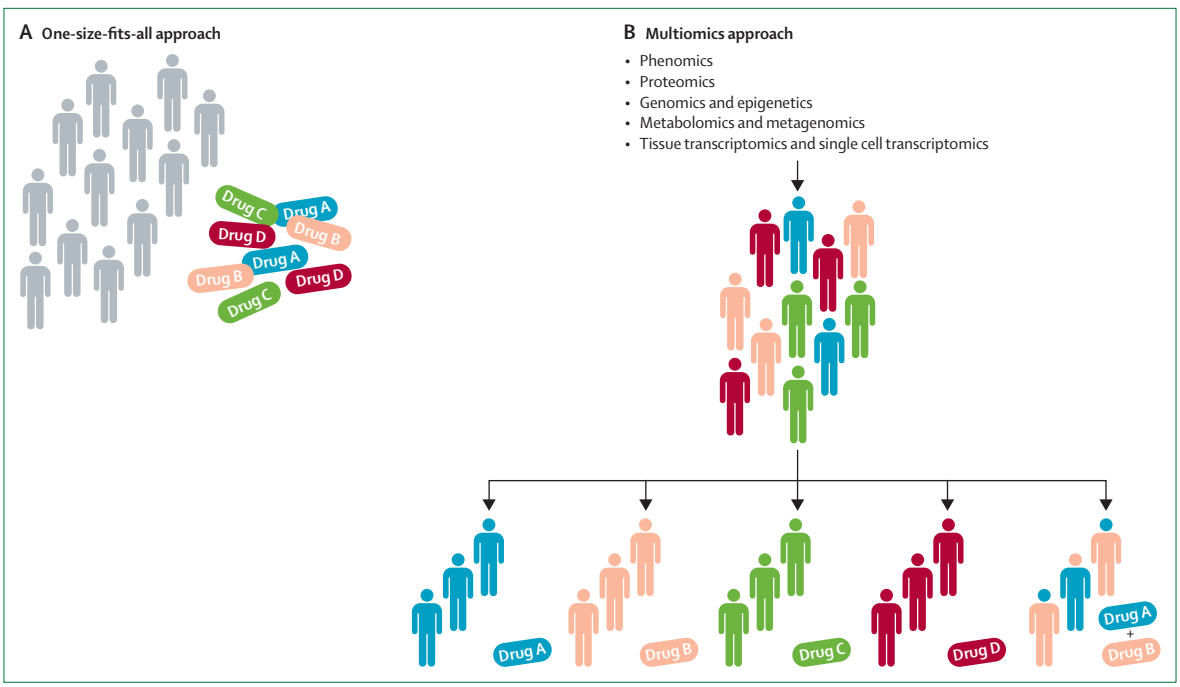

Figure 2: **Multiomics approach to personalised treatment in Crohn's disease**
A workflow showing the approach likely to be taken for future biomarker discovery, integrating a range of modern omics technologies to be used alongside accurate and extensive clinical phenotyping data. (A) Heterogenous population of patients with IBD characterised solely by clinical phenotypes and therapies initiated accordingly. (B) Heterogenous population of patients with IBD profiled by multiomics techniques and based on results given personalised IBD therapy according to their underlying biology, to ensure that patients receive the therapy, or therapies, that are most suitable for their disease. IBD=inflammatory bowel disease.

vedolizumab treatment. Combining clinical information with microbial taxa and pathway data allowed a model to be created to predict the likelihood of clinical remission at week 14. Crucially, a more diverse microbiome at baseline was associated with greater likelihood of vedolizumab response.[92] The reasons for this finding have yet to be determined but before speculating about possible mechanisms it will be important—as with other predictive markers—to ensure that this effect is validated in independent cohorts, and is actually reflective of underlying biology.

Anti-IL-12 and 23

IL-23 is a heterodimeric pro-inflammatory cytokine: one subunit, p40, is shared with IL-12 whereas the p19 subunit is unique to IL-23. Inhibition of the p40 subunit, using ustekinumab, has shown efficacy in treating Crohn's disease,[93] however real-world clinical use of this medication in Crohn's disease remains in the early stages. Similarly to previous medications in Crohn's disease, non-specific features have been associated with greater response to ustekinumab including no previous abdominal surgery and an uncomplicated disease phenotype.[94] However, this effect once again highlights the need to move away from retrospectively identified features that reflect a more severe past course of disease and a refocus on factors that might be useful in clinical practice for identifying individuals who will or will not respond to p40 or p19 inhibitors.

Transcriptomic and immune predictors of response

Given the success of p40 inhibition, subsequent focus has turned to p19 monoclonal antibody treatment and in a placebo-controlled trial of 119 patients with Crohn's disease, MEDI2070 (brazikumab) showed promising efficacy signals (49·2% in clinical remission at week 8).[95] Given the known role of IL-22 as an upstream regulator of IL-23 signalling, concentrations of IL-22 were measured at baseline and a concentration of 15·6 pg/mL or more was reported to be associated with an increased likelihood of clinical remission at week 8. These results were not statistically significant, but this finding might reflect a lack of power. Indeed, these results would be consistent with the EMBARK trial, a biomarker discovery trial of 107 patients with ulcerative colitis and 157 patients with Crohn's disease.[96] A subset of 66 patients with Crohn's disease in the EMBARK trial had concurrent ileocolonoscopy and CT enterography data, and using them as dependent variables, serum IL-22 was found to be associated with inflammatory disease activity. Although the relationship between serum and mucosal IL-22 concentrations warrants further investigation, these observations provide promise and encourage ongoing incorporation of biomarker discovery in early phase clinical trial design. Biomarkers developed using

such discovery approaches can then be tested for clinical utility in much larger and adequately powered late-phase clinical trials, as is anticipated for phase 3 trials of p19 antibody treatments.

Multiomic predictors of response

Multiomic applications incorporating some or all of the elements described above, rather than considering each in isolation, have recently been described and offer substantial promise.[97] Indeed, such a multiomics approach has been adopted to examine biological and endoscopic response in patients with Crohn's disease treated with ustekinumab.[98] Before initiation of ustekinumab therapy, CD14+ cells, CD4+ T cells, and inflamed colonic and ileal biopsies were collected and processed for RNA sequencing. Transcriptomic data were integrated with proteomic and genetic data using a multiomics factor analysis,[97] highlighting pathways associated with ustekinumab response and identified a ten-feature panel that predicted response to ustekinumab in the same cohort. Further work is underway to validate these omic signals in larger, independent cohorts. Crucially, this work should be supported by a greater multiomic study of large cohorts through collaborative efforts such as the IBD BioResource in the UK and IBD Multiomics Data project in the USA.[99,100] Such systems biology approaches hold substantial promise in facilitating biomarker discovery and could bring the prospect of personalised medicine closer to reality (figure 2, table 2).

Conclusions and future directions

There is much to be gained from using novel approaches to better understand the biology that determines the course of Crohn's disease and response to treatment. Indeed, greater understanding of the determinants of outcome at an individual patient level seems to be a prerequisite to achieve the goal of personalised medicine. In particular, accurate tests or biomarkers will be needed to help guide personalised therapy. Translating such biomarker studies to clinical benefit, however, presents many logistical challenges, including the need for careful study design and well phenotyped and prospectively collected patient cohorts. The incorporation of biomarker discovery into early-phase clinical trials should help facilitate this translation, and mean that potential biomarkers can be independently validated as part of phase 2 or 3 trials.

With increasing financial support for personalised medicine and substantial technological advances, prognostic and predictive biomarkers will continue to be developed from both academic and industrial efforts. As a result, it will be important for stringent standards to be applied to biomarker discovery, validation, and clinical translation. In particular, there should be a movement away from retrospective associations with a clinical phenotype or certainly greater scrutiny of retrospective associations to prove causation and not just correlation. Furthermore, to achieve these aims, there will probably need to be greater collaboration between academia and industry to allow an environment for such biomarker development and translation to flourish.

Although most current tools to predict prognosis and response to treatments are not yet ready for routine clinical use, advances on several fronts do appear close to clinical application (panel), and some are already available—thus providing substantial room for optimism. For example, the whole blood prognostic biomarker being investigated in the PROFILE trial is already available for clinical use (PredictSure IBD; PredictImmune, Cambridge. UK). Other tests are also in the early stages of being translated for clinical use, including genotyping *NUDT15* variants for risk of thiopurine-induced myelotoxicity, and specific HLA alleles for anti-TNF immunogenicity and thiopurine-induced pancreatitis.

Scientific advances, including the development of new technologies and novel computational approaches,

	Notes
Genetic	
PANTS (NCT03088449)	A prospective, uncontrolled cohort study of 1610 patients with IBD started on anti-TNF therapy; recruitment to this study has been completed and is now in the follow-up phase; aiming to provide novel insights into anti-TNF response and non-response
IBD BioResource	An observational study across the UK that is seeking to recruit 50 000 cases of IBD; aiming to further understand the functional effect of IBD-associated gene variants
Transcriptomic	
PROFILE (ISRCTN 11808228)	A biomarker-stratified trial in Crohn's disease seeking to recruit 400 patients to determine whether clinical outcomes can be optimised from diagnosis by using a blood-based prognostic biomarker to stratify therapy
RISK (NCT00790543)	An observational prospective study of treatment-naive, newly diagnosed paediatric Crohn's disease; 913 cases of Crohn's disease and 887 controls have been recruited and are currently undergoing clinical follow-up; seeking to determine novel genetic, transcriptomic, and microbial biomarkers associated with outcomes
Metabolomic or microbiomic	
PREdICCt (ISRCTN 11808228)	An observational study aiming to recruit 3100 patients with IBD in remission, seeking to determine environmental factors—including contributions from dietary intake and the gut microbiome—to both remission and relapse of inflammation
IBD Multiomics database	An in-depth multiomic profiling project of 90 participants over the course of 1 year with data then made publicly available to allow the scientific community to gain increased understanding of the complex interplay in IBD
Proteomic	
IBD-Character	A proteomic biomarker discovery study of 400 patients with newly diagnosed, treatment-naive IBD, 200 symptomatic patients without evidence of IBD, and 200 healthy age-matched controls; seeking to identify proteomic markers associated with clinical outcomes
PREDICTS	A biorepository platform study of military personnel from the USA, with 2000 IBD cases (1000 with Crohn's disease and 1000 with ulcerative colitis) and 500 healthy controls; seeking novel proteomic biomarkers particularly before development of IBD

IBD=inflammatory bowel disease. TNF=tumour necrosis factor.

Table 2: **Ongoing research studies likely to yield important findings for the field of personalised medicine in Crohn's disease**

Search strategy and selection criteria

We identified references for this Review through searches of PubMed using the search terms "Crohn's disease", "prognosis", "prediction", "response", "non-response", "biomarkers", "personalised medicine", and "precision medicine" from Jan 1, 1980, to July 31, 2019. Additional references were identified through reviewed articles. Only manuscripts published in English were reviewed. We searched abstracts presented at the annual meetings of the European Crohn's and Colitis Organisation, British Society of Gastroenterology, United European Gastroenterology, and Digestive Diseases Week. The final reference list was generated on the basis of originality and relevance to the broad scope of this Review.

coupled with an increased awareness of the importance of personalised medicine should facilitate the discovery of more prognostic and predictive biomarkers in the near future. Indeed, we anticipate that soon it will be possible to use a combination of molecular markers to help guide clinical decision making, as is current practice in oncology. This approach would enable clinicians to identify which patients require more intensive therapy and help them to select the drug that is most appropriate—thus making personalised medicine a reality in the treatment of Crohn's disease.

Contributors

NMN and BV did the literature search and review of identified articles. NMN and BV wrote the initial manuscript draft with input from MP and input and oversight from JCL. All authors approved the final version of the manuscript.

Declaration of interests

NMN, BV, and MP declare no competing interests. JCL has done consultancy work for PredictImmune.

Acknowledgments

NMN is supported by a Medical Research Council PhD Studentship (MC_UU_171339). BV is a doctoral fellow of the Research Foundation Flanders, Belgium (11Y9418N). JCL is supported by a Wellcome Trust Intermediate Clinical Fellowship (105920/Z/14/Z).

References

1 Burisch J, Jess T, Martinato M, Lakatos PL. The burden of inflammatory bowel disease in Europe. *J Crohns Colitis* 2013; **7:** 322–37.

2 Huang H, Fang M, Jostins L, et al. Fine-mapping inflammatory bowel disease loci to single-variant resolution. *Nature* 2017; **547:** 173–78.

3 Kaser A, Zeissig S, Blumberg RS. Inflammatory bowel disease. *Annu Rev Immunol* 2010; **28:** 573–621.

4 Pascal V, Pozuelo M, Borruel N, et al. A microbial signature for Crohn's disease. *Gut* 2017; **66:** 813–22.

5 Aden K, Rehman A, Waschina S, et al. Metabolic functions of gut microbes associate with efficacy of tumor necrosis factor antagonists in patients with inflammatory bowel diseases. *Gastroenterology* 2019; published online July 18. DOI:10.1053/j.gastro.2019.07.025.

6 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**: 660–67.

7 Khanna R, Bressler B, Levesque BG, et al. Early combined immunosuppression for the management of Crohn's disease (REACT): a cluster randomised controlled trial. *Lancet* 2015; **386:** 1825–34.

8 Hart AL, Lomer M, Verjee A, et al. What are the top 10 research questions in the treatment of inflammatory bowel disease? *J Crohn's Colitis* 2017; **11:** 204–11.

9 Cameron D, Piccart-Gebhart MJ, Gelber RD, et al. 11 years' follow-up of trastuzumab after adjuvant chemotherapy in HER2-positive early breast cancer: final analysis of the HERceptin Adjuvant (HERA) trial. *Lancet* 2017; **389**: 1195–205.

10 Sparano JA, Gray RJ, Ravdin PM, et al. Clinical and genomic risk to guide the use of adjuvant therapy for breast cancer. *N Engl J Med* 2019; **380:** 2395–405.

11 Collins FS, Varmus H. A new initiative on precision medicine. *N Engl J Med* 2015; **372:** 793–95.

12 McNamee CJ. UK pharmacogenetics and stratified medicine network. *Per Med* 2016; **13:** 107–12.

13 Beaugerie L, Seksik P, Nion-Larmurier I, Gendre JP, Cosnes J. Predictors of Crohn's disease. *Gastroenterology* 2006; **130:** 650–56.

14 Torres J, Caprioli F, Katsanos KH, et al. Predicting outcomes to optimize disease management in inflammatory bowel diseases. *J Crohn's Colitis* 2016; **10:** 1385–94.

15 Cosnes J, Bourrier A, Nion-Larmurier I, Sokol H, Beaugerie L, Seksik P. Factors affecting outcomes in Crohn's disease over 15 years. *Gut* 2012; **61:** 1140–45.

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

17 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.

18 Lee JC, Biasci D, Roberts R, et al. Genome-wide association study identifies distinct genetic contributions to prognosis and susceptibility in Crohn's disease. *Nat Genet* 2017; **49:** 262–68.

19 Walker GJ, Harrison JW, Heap GA, et al. Association of genetic variants in *NUDT15* with thiopurine-induced myelosuppression in patients with inflammatory bowel disease. *JAMA* 2019; **321:** 773–85.

20 Yang SK, Hong M, Baek J, et al. A common missense variant in *NUDT15* confers susceptibility to thiopurine-induced leukopenia. *Nat Genet* 2014; **46:** 1017–20.

21 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:** 1131–34.

22 Sazonovs A, Kennedy N, Moutsianas L, et al. *HLA-DQA1*05* is associated with the development of antibodies to anti-TNF therapy. *bioRxiv* 2018; published online Sept 9. DOI:10.1101/410035 (preprint).

23 Heliö T, Halme L, Lappalainen M, et al. *CARD15/NOD2* gene variants are associated with familially occurring and complicated forms of Crohn's disease. *Gut* 2003; **52:** 558–62.

24 Siegel CA, Siegel LS, Hyams JS, et al. Real-time tool to display the predicted disease course and treatment response for children with Crohn's disease. *Inflamm Bowel Dis* 2011; **17:** 30–38.

25 Siegel CA, Horton H, Siegel LS, et al. A validated web-based tool to display individualised Crohn's disease predicted outcomes based on clinical, serologic and genetic variables. *Aliment Pharmacol Ther* 2016; **43:** 262–71.

26 Dubinsky MC, Kugathasan S, Kwon S, et al. Multidimensional prognostic risk assessment identifies association between *IL12B* variation and surgery in Crohn's disease. *Inflamm Bowel Dis* 2013; **19:** 1662–70.

27 Verstockt B, Cleynen I. Genetic influences on the development of fibrosis in Crohn's disease. *Front Med* 2016; **3:** 24.

28 Jung C, Colombel JF, Lemann M, et al. Genotype/phenotype analyses for 53 Crohn's disease associated genetic polymorphisms. *PLoS One* 2012; **7:** e52223.

29 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:** 388–95.

30 Ananthakrishnan AN, Huang H, Nguyen DD, Sauk J, Yajnik V, Xavier RJ. Differential effect of genetic burden on disease phenotypes in Crohn's disease and ulcerative colitis: analysis of a north American cohort. *Am J Gastroenterol* 2014; **109:** 395–400.

31 Lee JC, Espéli M, Anderson CA, et al. Human SNP links differential outcomes in inflammatory and infectious disease to a *FOXO3*-regulated pathway. *Cell* 2013; **155:** 57–69.

32 Heap GA, So K, Weedon M, et al. Clinical features and HLA association of 5-aminosalicylate (5-ASA)-induced nephrotoxicity in inflammatory bowel disease. *J Crohn's Colitis* 2016; **10:** 149–58.

33 Kennedy NA, Heap GA, Green HD, et al. Predictors of anti-TNF treatment failure in anti-TNF-naive patients with active luminal Crohn's disease: a prospective, multicentre, cohort study. *Lancet Gastroenterol Hepatol* 2019; **4:** 341–53.

34 Kugathasan S, Denson LA, Walters TD, et al. Prediction of complicated disease course for children newly diagnosed with Crohn's disease: a multicentre inception cohort study. *Lancet* 2017; **389:** 1710–18.

35 Haberman Y, Tickle TL, Dexheimer PJ, et al. Pediatric Crohn disease patients exhibit specific ileal transcriptome and microbiome signature. *J Clin Invest* 2014; **124:** 3617–33.

36 Marigorta UM, Denson LA, Hyams JS, et al. Transcriptional risk scores link GWAS to eQTLs and predict complications in Crohn's disease. *Nat Genet* 2017; **49:** 1517–21.

37 Telesco SE, Brodmerkel C, Zhang H, et al. Gene expression signature for prediction of golimumab response in a phase 2a open-label trial of patients with ulcerative colitis. *Gastroenterology* 2018; **155:** 1008–11.

38 Shen-Orr SS, Gaujoux R. Computational deconvolution: extracting cell type-specific information from heterogeneous samples. *Curr Opin Immunol* 2013; **25:** 571–78.

39 McKinney EF, Lyons PA, Carr EJ, et al. A CD8+T cell transcription signature predicts prognosis in autoimmune disease. *Nat Med* 2010; **16:** 586–91.

40 McKinney EF, Lee JC, Jayne DRW, Lyons PA, Smith KGC. T-cell exhaustion, co-stimulation and clinical outcome in autoimmunity and infection. *Nature* 2015; **523:** 612–16.

41 Allez M, Auzolle C, Ngollo M, et al. T cell clonal expansions in ileal Crohn's disease are associated with smoking behaviour and postoperative recurrence. *Gut* 2019; published online Feb 12. DOI:10.1136/gutjnl-2018-317878.

42 Biasci D, Lee JC, Noor NM, et al. A blood-based prognostic biomarker in IBD. *Gut* 2019; **68:** 1386–95.

43 Parkes M, Noor NM, Dowling F, et al. PRedicting Outcomes For Crohn's dIsease using a moLecular biomarkEr (PROFILE): protocol for a multicentre, randomised, biomarker-stratified trial. *BMJ Open* 2018; **8:** e026767.

44 Choung RS, Princen F, Stockfisch TP, et al. Serologic microbial associated markers can predict Crohn's disease behaviour years before disease diagnosis. *Aliment Pharmacol Ther* 2016; **43:** 1300–10.

45 Zhang Z, Li C, Zhao X, et al. Anti-*Saccharomyces cerevisiae* antibodies associate with phenotypes and higher risk for surgery in Crohn's disease: a meta-analysis. *Dig Dis Sci* 2012; **57:** 2944–54.

46 Ferrante M, Henckaerts L, Joossens M, et al. New serological markers in inflammatory bowel disease are associated with complicated disease behaviour. *Gut* 2007; **56:** 1394–403.

47 Maaser C, Sturm A, Vavricka SR, et al. ECCO-ESGAR Guideline for Diagnostic Assessment in IBD part 1: initial diagnosis, monitoring of known IBD, detections of complications. *J Crohns Colitis* 2019; **13:** 144–64.

48 Vandeputte D, Kathagen G, D'Hoe K, et al. Quantitative microbiome profiling links gut community variation to microbial load. *Nature* 2017; **551:** 507–11.

49 Sokol H, Pigneur B, Watterlot L, et al. *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proc Natl Acad Sci* 2008; **105:** 16731–36.

50 Britton GJ, Contijoch EJ, Mogno I, et al. Microbiotas from humans with inflammatory bowel disease alter the balance of gut Th17 and RORγt+ regulatory t cells and exacerbate colitis in mice. *Immunity* 2019; **50:** 212–24.

51 Biedermann L, Zeitz J, Mwinyi J, et al. Smoking cessation induces profound changes in the composition of the intestinal microbiota in humans. *PLoS One* 2013; **8:** e59260.

52 David LA, Maurice CF, Carmody RN, et al. Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 2014; **505:** 559–63.

53 Henriksen M, Jahnsen J, Lygren I, et al. C-reactive protein: a predictive factor and marker of inflammation in inflammatory bowel disease. Results from a prospective population-based study. *Gut* 2008; **57:** 1518–23.

54 Oh K, Oh EH, Baek S, et al. Elevated C-reactive protein level during clinical remission can predict poor outcomes in patients with Crohn's disease. *PLoS One* 2017; **12:** e0179266.

55 Colombel JF, Panaccione R, Bossuyt P, et al. Effect of tight control management on Crohn's disease (CALM): a multicentre, randomised, controlled phase 3 trial. *Lancet* 2018; **390:** 2779–89.

56 De Vos M, Dewit O, D'Haens G, et al. Fast and sharp decrease in calprotectin predicts remission by infliximab in anti-TNF naïve patients with ulcerative colitis. *J Crohn's Colitis* 2012; **6:** 557–62.

57 De Vos M, Louis EJ, Jahnsen J, et al. Consecutive fecal calprotectin measurements to predict relapse in patients with ulcerative colitis receiving infliximab maintenance therapy. *Inflamm Bowel Dis* 2013; **19:** 2111–17.

58 Wright EK, Kamm MA, De Cruz P, et al. Measurement of fecal calprotectin improves monitoring and detection of recurrence of Crohn's disease after surgery. *Gastroenterology* 2015; **148:** 938–47.

59 Brookes MJ, Whitehead S, Gaya DR, Hawthorne AB. Practical guidance on the use of faecal calprotectin. *Frontline Gastroenterol* 2017; **9:** 87–91.

60 Gecse KB, Brandse JF, Van Wilpe S, et al. Impact of disease location on fecal calprotectin levels in Crohn's disease. *Scand J Gastroenterol* 2015; **50:** 841–47.

61 Kalla R, Kennedy NA, Hjelm FN, et al. Proximity extension assay technology identifies novel serum biomarkers for predicting inflammatory bowel disease: IBD Character Consortium. *Gastroenterology* 2015; **148:** S-436.

62 Porter CK, Riddle MS, Gutierrez RL, et al. Cohort profile of the PRoteomic Evaluation and Discovery in an IBD Cohort of Tri-service Subjects (PREDICTS) study: rationale, organization, design, and baseline characteristics. *Contemp Clin Trials Commun* 2019; **14:** 100345.

63 Chevaux JB, Peyrin-Biroulet L, Sparrow MP. Optimizing thiopurine therapy in inflammatory bowel disease. *Inflamm Bowel Dis* 2011; **17:** 1428–35.

64 Ben-Horin S, Chowers Y. Tailoring anti-TNF therapy in IBD: drug levels and disease activity. *Nat Rev Gastroenterol Hepatol* 2014; **11:** 243–55.

65 D'Haens G, Vermeire S, Lambrecht G, et al. Increasing infliximab dose based on symptoms, biomarkers, and serum drug concentrations does not increase clinical, endoscopic, and corticosteroid-free remission in patients with active luminal crohn's disease. *Gastroenterology* 2018; **154:** 1343–51.

66 Parsi MA, Achkar J, Richardson S, et al. Predictors of response to infliximab in patients with Crohn's disease. *Gastroenterology* 2002; **123:** 707–13.

67 Arnott IDR, McNeill G, Satsangi J. An analysis of factors influencing short-term and sustained response to infliximab treatment for Crohn's disease. *Aliment Pharmacol Ther* 2003; **17:** 1451–57.

68 Louis E, Vermeire S, Rutgeerts P, et al. A positive response to infliximab in Crohn disease: association with a higher systemic inflammation before treatment but not with –308 TNF gene polymorphism. *Scand J Gastroenterol* 2002; **37:** 818–24.

69 Sandborn WJ, Colombel JF, D'Haens G, et al. Association of baseline C-reactive protein and prior anti-tumor necrosis factor therapy with need for weekly dosing during maintenance therapy with adalimumab in patients with moderate to severe Crohn's disease. *Curr Med Res Opin* 2013; **29:** 483–93.

70 Vermeire S, Louis E, Carbonez A, et al. Demographic and clinical parameters influencing the short-term outcome of anti-tumor necrosis factor (infliximab) treatment in Crohn's disease. *Am J Gastroenterol* 2002; **97:** 2357–63.

71 Yoon SM, Haritunians T, Chhina S, et al. Colonic phenotypes are associated with poorer response to anti-TNF therapies in patients with IBD. *Inflamm Bowel Dis* 2017; **23:** 1382–93.

72 Atreya R, Neumann H, Neufert C, et al. In vivo imaging using fluorescent antibodies to tumor necrosis factor predicts therapeutic response in Crohn's disease. *Nat Med* 2014; **20:** 313–18.

73 Barber GE, Yajnik V, Khalili H, et al. Genetic markers predict primary non-response and durable response to anti-TNF biologic therapies in Crohn's disease. *Am J Gastroenterol* 2016; **111:** 1816–22.

74 Billiet T, Papamichael K, de Bruyn M, et al. A matrix-based model predicts primary response to infliximab in Crohn's disease. *J Crohns Colitis* 2015; **9:** 1120–26.

75 Arijs I, Li K, Toedter G, et al. Mucosal gene signatures to predict response to infliximab in patients with ulcerative colitis. *Gut* 2009; **58:** 1612–19.

76 Arijs I, Quintens R, Van Lommel L, et al. Predictive value of epithelial gene expression profiles for response to infliximab in Crohn's disease. *Inflamm Bowel Dis* 2010; **16:** 2090–98.

77 Weiser M, Simon JM, Kochar B, et al. Molecular classification of Crohn's disease reveals two clinically relevant subtypes. *Gut* 2018; **67:** 36–42.

78 West NR, Hegazy AN, Owens BMJ, et al. Oncostatin M drives intestinal inflammation and predicts response to tumor necrosis factor-neutralizing therapy in patients with inflammatory bowel disease. *Nat Med* 2017; **23:** 579–89.

79 Billiet T, Cleynen I, Ballet V, et al. Evolution of cytokines and inflammatory biomarkers during infliximab induction therapy and the impact of inflammatory burden on primary response in patients with Crohn's disease. *Scand J Gastroenterol* 2017; **52:** 1086–92.

80 Verstockt B, Verstockt S, Dehairs J, et al. Low TREM1 expression in whole blood predicts anti-TNF response in inflammatory bowel disease. *EBioMedicine* 2019; **40:** 733–42.

81 Gaujoux R, Starosvetsky E, Maimon N, et al. Cell-centred meta-analysis reveals baseline predictors of anti-TNFα non-response in biopsy and blood of patients with IBD. *Gut* 2019; **68:** 604–14.

82 Verstockt B, Verstockt S, Blevi H, et al. TREM-1, the ideal predictive biomarker for endoscopic healing in anti-TNF-treated Crohn's disease patients? *Gut* 2019; **68:** 1531–33.

83 Martin JC, Boschetti G, Chang C, et al. Single-cell analysis of Crohn's disease lesions identifies a pathogenic cellular module associated with resistance to anti TNF therapy. *bioRxiv* 2018; published online Dec 20. DOI:10.1101/503102 (preprint).

84 Kolho KL, Korpela K, Jaakkola T, et al. Fecal microbiota in pediatric inflammatory bowel disease and its relation to inflammation. *Am J Gastroenterol* 2015; **110:** 921–30.

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

86 Feagan BG, Rubin DT, Danese S, et al. Efficacy of vedolizumab induction and maintenance therapy in patients with ulcerative colitis, regardless of prior exposure to tumor necrosis factor antagonists. *Clin Gastroenterol Hepatol* 2017; **15:** 229–39.

87 Dulai PS, Boland BS, Singh S, et al. Development and validation of a scoring system to predict outcomes of vedolizumab treatment in patients with Crohn's disease. *Gastroenterology* 2018; **155:** 687–95.

88 Rath T, Bojarski C, Neurath MF, Atreya R. Molecular imaging of mucosal α4β7 integrin expression with the fluorescent anti-adhesion antibody vedolizumab in Crohn's disease. *Gastrointest Endosc* 2017; **86:** 406–08.

89 Vermeire S, O'Byrne S, Keir M, et al. Etrolizumab as induction therapy for ulcerative colitis: a randomised, controlled, phase 2 trial. *Lancet* 2014; **384:** 309–18.

90 Tew GW, Hackney JA, Gibbons D, et al. Association between response to etrolizumab and expression of integrin αe and granzyme a in colon biopsies of patients with ulcerative colitis. *Gastroenterology* 2016; **150:** 477–87.

91 Arijs I, De Hertogh G, Lemmens B, et al. Effect of vedolizumab (anti-α4β7-integrin) therapy on histological healing and mucosal gene expression in patients with UC. *Gut* 2018; **67:** 43–52.

92 Ananthakrishnan AN, Luo C, Yajnik V, et al. Gut microbiome function predicts response to anti-integrin biologic therapy in inflammatory bowel diseases. *Cell Host Microbe* 2017; **21:** 603–10.

93 Feagan BG, Sandborn WJ, Gasink C, et al. Ustekinumab as induction and maintenance therapy for Crohn's disease. *N Engl J Med* 2016; **375:** 1946–60.

94 Barré A, Colombel JF, Ungaro R. Review article: predictors of response to vedolizumab and ustekinumab in inflammatory bowel disease. *Aliment Pharmacol Ther* 2018; **47:** 896–905.

95 Sands BE, Chen J, Feagan BG, et al. Efficacy and safety of MEDI2070, an antibody against interleukin 23, in patients with moderate to severe Crohn's disease: a phase 2a study. *Gastroenterology* 2017; **153:** 77–86.

96 Faubion WA, Fletcher JG, O'Byrne S, et al. Emerging biomarkers in inflammatory bowel disease (EMBARK) study identifies fecal calprotectin, serum MMP9, and serum IL-22 as a novel combination of biomarkers for Crohn's disease activity: role of cross-sectional imaging. *Am J Gastroenterol* 2013; **108:** 1891–900.

97 Argelaguet R, Velten B, Arnol D, et al. Multi-omics factor analysis—a framework for unsupervised integration of multi-omics data sets. *Mol Syst Biol* 2018; **14:** e8124.

98 Verstockt B, Sudahakar P, Creyns B, et al. An integrated multi-omics biomarker predicting endoscopic response in ustekinumab treated patients with Crohn's disease. *J Crohns Colitis* 2019; **13:** S072–73.

99 Parkes M, IBD BioResource Investigators. IBD BioResource: an open-access platform of 25 000 patients to accelerate research in Crohn's and Colitis. *Gut* 2019; **68:** 1537–40.

100 Lloyd-Price J, Arze C, Ananthakrishnan AN, et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* 2019; **569:** 655–62.